School of GeoSciences

Dissertation
For the degree of

MSc in Geographical Information Science

NABILAH NAHARUDIN

August 2014

APPLICATION OF

LOCATION/ALLOCATION MODELS AND GIS

TO THE LOCATION OF NATIONAL PRIMARY SCHOOLS

IN RAWANG, MALAYSIA

By Nabilah Naharudin

Dissertation submitted as a part of
MSc in Geographical Information Science
August 2014
Institute of Geography
The University of Edinburgh

COPYRIGHT

STATEMENT OF ORIGINALITY

I declare that this dissertation represents my own work, and that where the work of others has been used it has been duly accredited. I further declare that the length of the components of this dissertation is   words for the Research Paper and   words for the Technical Report.

Signed:________________________________ Date:__________________________________

ACKNOWLEDGEMENT

Thanks to the Almighty. I would never have been able to finish my dissertation without His blessing and will.

I would like to express my gratitude to my supervisor, Dr. Neil Stuart, and also to Dr. Robert L. Hodgart for their continuous guidance throughout this research. Their supportive comments and constructive criticism greatly stimulated me and helped to carry me forward to finish this dissertation.

I also want to thank the Department of Education of the District of Gombak, the Department of Town and Regional Planning of the State of Selangor and the Department of Statistics Malaysia for providing much valuable information needed for this dissertation.

Special thanks to my financial sponsor, Majlis Amanah Rakyat (MARA), for giving me this huge opportunity to study here.

To all my closest friends (you know who you are), thank you so much. I could never be who I am without you.

Last but not least, to my parents and brothers, I can’t express how grateful I am to have all of you by my side ever since I was born.

Thank you.

PART I

RESEARCH PAPER

ABSTRACT

As education is compulsory in Malaysia, all children aged 7 to 12 are expected to attend primary schools. It is therefore a challenge for the government to provide schools which are well located to serve the children conveniently. Various criteria for siting schools may be assessed by using location/ allocation models (LAMs), a widely used tool for finding good locations for public facilities. The key factor in siting schools is the location of demand or number of children that will attend the schools involved, here the national primary schools of Malaysia. Based on census data and statistics reported in the Malaysia Education Blueprint, the estimated number of pupils from 23 neighbourhood areas likely to attend national primary schools was estimated using information on the age and ethnic breakdown of population in these units. Then, to increase accuracy in estimating the distribution of demand, estimated numbers of pupils from the 23 neighbourhoods were assigned at a finer scale to 10,718 cadastral plot units which gave quite precise locations for where pupils live. This distribution of demand to a fine scale then helped in assessing how well schools were located to serve the study area concerned using LAMs. First, the current distribution of schools and their catchment areas were examined. Then, various possibilities for improving accessibility to schools were explored by using 2 scenarios involving adding 1 or 2 new schools and using three different types of LAM with and without capacity constraints, including also one scenario involving closing 1 of the 5 existing schools which seemed less well located. Finally, the more realistic solutions found were selected for further consideration and evaluation as a basis for enhancing the accessibility of pupils to national primary schools in Rawang.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DOSM	Department of Statistics, Malaysia
DTRP	Department of Town and Regional Planning
DTRPS	Department of Town and Regional Planning Selangor
EB	Enumeration Block
GIS	Geographical Information System
LA	Location-Allocation
LAM	Location-Allocation Modelling
LAMs	Location-Allocation Models
LAP	Location-Allocation Problem
MOE	Ministry of Education
SRCPU	Small Residential Cadastral Plot Unit

1. INTRODUCTION

Public facilities for education such as primary and secondary schools are among the most important public service amenities as all children need to attend them. Thus, they should be sited by both planners and policy makers at locations with good accessibility to users. The standard planning policy in Malaysia for building such facilities appears to take account of a number of factors, including the required area of space, the type of land and also the population of the surrounding area according to the guidelines of the Department of Town and Regional Planning of Selangor (DTRPS) in 2010. The key factor is, however, the population of children in the surrounding area because schools are primarily designed to serve demand derived from population nearby.

Location-allocation models (LAMs) which attempt to find good locations for service facilities can be very useful tools for planning the locations of schools (Yeh & Chow, 1996). LAMs attempt to find accessible locations for one or more facilities so that, over the whole region concerned, the population served enjoys the best level of accessibility possible or something close to that. They provide spatial decision support systems that give sets of solutions on which decision makers can focus their discussions. This dissertation is mainly concerned with applying LAMs, integrated with a Geographical Information System (GIS), to locate schools, using travel distance through the local road network as the measure of accessibility.

If the results from LAMs are to be as helpful to planners as possible, it is essential to use accurate and detailed information on where children live and the routes they are likely to travel through the road network to school. Information on the location of population and the relevant age groups has to be obtained from the 2010 census carried out by the Federal Department of Statistics Malaysia (DOSM). The smallest areas for which such population data are available may contain 500-600 people or even some 1500 and can extend over many streets. A major concern of the present dissertation is therefore to use such data to estimate more precisely where children actually live by distributing these areal totals to the houses and apartments where children are likely to reside, using information on cadastral lots and a detailed street network.

More specifically, this dissertation explores various solutions to the problem of locating schools and allocating children to them in the city of Rawang, Selangor State, Malaysia. LAMs are used to examine the current locations of the schools to see how well they meet the Federal Government's goal of all children being within 800 meters from school and to help identify any poorly served areas (DTRPS, 2010). Then, various possible solutions for improving accessibility to badly served areas will be assessed by using LAMs which allow the provision of one or two new schools in addition to the existing 5 schools and which also explore the consequences of closing any apparently poorly located schools.

1.1. Schools in Malaysia

Primary schools in Malaysia are attended by children aged seven to twelve. Parents can choose to enrol their children in either public or private schools. Public primary schools are divided into two types, both free: national primary schools which use Malay as the first language; and vernacular primary schools which use Chinese, Indian or English as the medium of instruction. All public schools are under the responsibility of the Federal Ministry of Education (MOE). As primary education in Malaysia is compulsory, the government has been working hard to ensure all children aged seven to twelve are able to attend primary school. Therefore, it is essential for schools to be accessible to children. Unless otherwise stated, for convenience, here we will simply use 'school' to mean a national primary school.

The planning of schools is under each State's Department of Town and Regional Planning (DTRP). There are a number of planning criteria which are supposed to serve as guidelines for school location and for allocation of pupils from catchment areas to schools to ensure demand is well served. The two main criteria are as follows.

i. The population to be served must be within the 3000 – 7500 range for a school to be built in an area. However, ethnic and demographic composition also have to be taken into consideration, because within this population range, there may not be enough Malay children to justify a national school or enough Chinese or Indian children to justify a vernacular school for either group. Thus, the ethnic character of the population can be crucial for planning primary schools.

ii. Ideally, no pupil should be more than 0.8 kilometres (800 metres) from school or travel more than 10 minutes.

The process of allocating pupils to a primary school is a responsibility of the Federal MOE. MOE takes the number of classrooms available in that particular school into consideration in this process with the recommended maximum capacity of any school based on a figure of 30 pupils per classroom. Thus, if there are 2 classrooms in a school for each of the 6 age groups and each of these 12 classes has 30 pupils, the capacity of the school will be 360 pupils. Normally, pupils are supposed to be assigned to the nearest school to their home. If the closest school has already achieved maximum capacity, however, pupils will normally be assigned to the second nearest school and so on (Abdul Rahman, 2014).

2. LOCATION-ALLOCATION MODELS

Location-allocation modelling (LAM) is the process of determining 'optimal' locations for one or more facilities so that the area's population is served in the most convenient manner possible (Schietzelt & Densham, 2003). These models optimize efficiency by simultaneously determining the locational configurations of the facilities and assigning population to facilities (allocation), hence the term location-allocation (Turner, 2006; Chang, 2012)

2.1. Main Types of Location Allocation Models

In a review of the wide range of location-allocation models which had developed up to that point, Hodgart (1978) noted that they can be classified into groups or families according to:

(a) whether they use the geometrical frameworks of planes or networks (the latter are usually more realistic);

(b) whether they prioritise objectives based on spatial efficiency or spatial equity and equality;

(c) the specific goals used to measure spatial efficiency or equity;

(d) whether demand is assigned to the nearest centre or whether catchment areas for facilities may be allowed to overlap to some degree.

A further distinction could be between models with and without capacity constraints on facilities. The present study will only be concerned with network space. As regards objectives, models which prioritise spatial efficiency through minimising aggregate travel distance in the system (often called p-median models) will be used as well as models which can strike a flexible balance between efficiency and equity through attempting to maximise the population 'covered' i.e. within the range of a covering radius, which Hodgart (1978) notes can be seen as a mixture of efficiency and equity. Models both with and without capacity constraints will be used here, as schools can be seen as ultimately having limits to their capacity, even if these are not precise. When capacity constraints are used, some pupils may be allocated to their second or even third nearest school, so catchment areas may not be based on nearest centre assignments. Models where allocation is done through use of spatial interaction models to produce catchment areas which overlap to varying degrees, as described by Hodgart (1978 and 1981), are not used.

It may be worth noting here that minimizing the total travel time or total distance over the whole system is mathematically equivalent to minimizing mean time or mean travel distance. LAMs can be used in the private sector to locate warehouses and other private facilities. In the public sector, they can be used widely in locating emergency services, schools, health centres and other public facilities (Blazevic, 2006). To solve such problems, one possible objective for schools is to maximize the number of pupils within a desirable or required travel distance from school (termed radius of cover), known as the maximal covering location problem (MCLP).

2.2. Location-Allocation Problems for Schools

The usage of location-allocation models for school locations has been discussed by several authors (Møller-Jensen, 1998; Ariffin, 2011; Handayani, 1997; Pearce, 2000; London, 2007). The allocation stage typically involves the process of identifying specific areal units or small districts in which demand is located and assigning each of these to one or more schools (usually all the node's demand is assigned to the nearest facility with free capacity). Thus many location-allocation models may require four key components: facilities, the capacities of the facilities (where appropriate), a measure of demand and a measure of 'impedance' such as distance or time travelled.

A network model essentially consists of a set of connections or routes between a set of nodes made through links (or 'arcs') which meet at the nodes i.e. a topological structure (Goodchild, 1992). In this study, a network model based on detailed information on all roads in the study area, including minor roads, will be used to find the shortest travel paths through the road network for pupils travelling from the 10,718 small residential areas, represented as demand nodes, to schools treated as supply nodes and thereby estimate the 'impedance' to be overcome in travelling from home to school. This network model consists of 3,956 links and 1,426 nodes and can be regarded as giving a fairly realistic picture of where pupils live and how they travel to school.

LAMs using a network reflect how the transportation of resources on the ground normally takes place along certain pre-defined corridors (Lwin & Murayama, 2012). Each linear feature or link in the network has an impedance value indicating the cost of moving along that line. Here we will only use travel distance.

Implementing the location-allocation models for schools requires digital data with sufficient spatial detail to describe and model the properties of the street network and the distribution of demand accurately. (Handayani, 1997). The number of pupils allocated to a particular school and the corresponding range of distances pupils need to travel can then be reviewed.

2.3. Study Area

Rawang is located about 20 kilometres north-west of Kuala Lumpur in the district Gombak in the State of Selangor (Figure 1). The town is currently experiencing rapid development due to its close proximity to Kuala Lumpur. The constant flow of construction vehicles in and out of the town indicates that much of the urban area is now undergoing transformation. Based on statistics from DOSM, in 2010 there were 38, 416 people living in the town. The Local Plan for Rawang-Kuang states that, for every 6, 250 people living in an area, the government should provide a school. Currently, however, there are only five schools in Rawang to serve its expanding population.

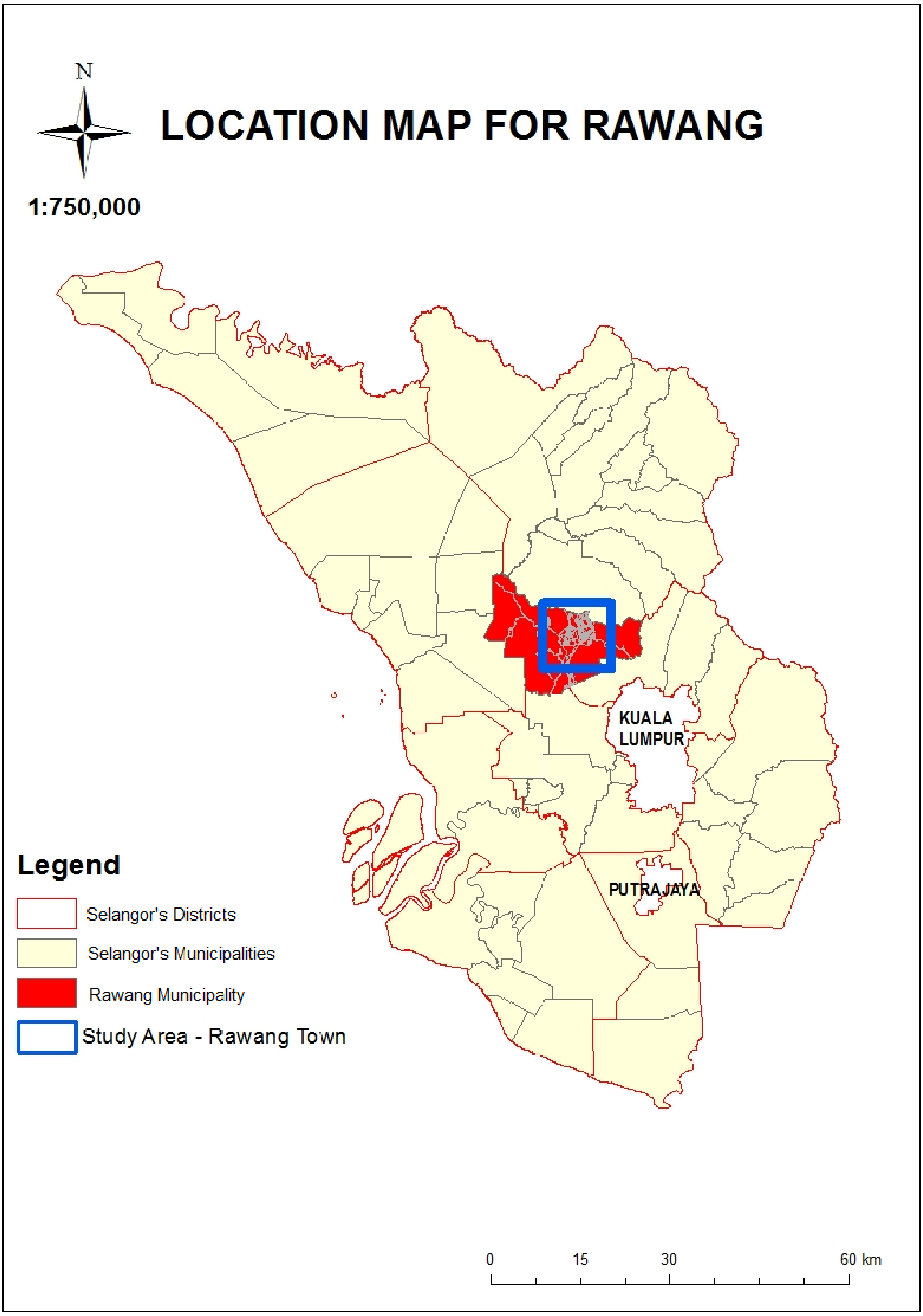

Figure 1 Location Map for Rawang

3. METHODOLOGY

This section will discuss the particular range of Location-Allocation Models used and their implementation in ArcGIS. Finally, the way census data and cadastral data are used to estimate the home locations of pupils is briefly outlined.

3.1. Location-Allocation Models in ArcGIS

For this study, the ArcGIS software (Version 10.1) is used because, within the range of 7 LAMs it supports several which are appropriate for school location problems are now included and its GIS components facilitate display of the results (ESRI, 2012). These seven common LA problems are:

i. Maximize Coverage i.e. the number of service users 'covered' i.e. within a specified distance of a facility (MCLP model);
ii. Minimize Impedance i.e. minimize aggregate travel distance or time or cost of users;
iii. Maximize Capacitated Coverage;
iv. Minimize Facilities;
v. Maximize Attendance;
vi. Maximize Market Share;
vii. Target Market Share.

Only the first four models will be used here. The last two problems are not relevant as they involve location problems where private supply centres are competing in a market, which is not appropriate for schools.

The fifth problem is that of maximizing the use of or attendance at facilities when the level of attendance declines with distance from supply, which is not appropriate for schools. To ensure that as many primary school pupils as possible are within specified or desirable distances of a school and to reduce the overall distance pupils travel to school, the Maximize Coverage and Minimize Impedance problems respectively are clearly very relevant here. Additionally, the Maximum Capacitated

Coverage problem is also particularly relevant because in maximizing cover it also takes account of the capacity of each school. The Minimize Facilities problem may also be helpful because it finds the smallest number of facilities needed to ensure all pupils enjoy a specified level of cover.

3.1.1.Maximizing Cover

As implemented in ArcGIS 10.1, the maximizing cover model allocates demand to facilities so that:

i. any demand will be considered as uncovered and unallocated if it is located outside the 'impedance cutoff' of all facilities;
ii. demand points within the 'impedance cutoff' of only one facility will have all their demand allocated to that facility;
iii. a demand point located within the 'impedance cutoff' of more than one facility will be assigned to the closest.

The Maximize Coverage problem has been formulated as follows by Marianov and Serra (2002):

$$\text{Maximize } \left\{ z = \sum_{i \in I} a_i y_i \right\} \tag{1}$$

Subject to:

$$y_i \leq \sum_{j \in N_i} x_j \ \forall i \ \in I \tag{2}$$

$$\sum_{j \in J} x_j = p \tag{3}$$

$$x_j, y_i = (0,1) \quad \forall j \in J, i \in I$$

The relevant notation is defined as follows:

I = Set of demand locations,

J = Set of candidate facility sites

P = The number of facilities to be located

x_j = 1 if facility is sited at j , 0 otherwise

y_i = 1 if demand node i is covered, 0 otherwise

S = Standard time or distance for coverage

N_i = The set of all candidate sites which can cover demand nodes i

a_i = The population or demand at demand node i

Equation (1) maximizes the amount of demand covered. Equation (2) states that the demand at node i is covered whenever at least one facility is located within the standard time or distance S. Equation (3) specifies the total number of facilities to be located.

3.1.2.Minimizing Impedance i.e. Minimizing Aggregate Travel Distance etc. (the p-median)

This problem, as addressed by Hakimi (1969), aims to minimize the distance travelled by the population to facilities, thereby maximising efficiency of access. The p-median problem can also include a distance or impedance cut off constraint such that any population beyond this maximum distance is not served. In the latter model, the facilities will be located at the most accessible points for serving the population within what can be regarded as a manageable distance or within an extreme limit for feasible travel.

The p-median model has various applications in both public and private sectors as it is always better that the total travel distance of pupils etc. is low so that users have better accessibility. The formulation of the p-median problem by Marianov and Serra (2002) is as follows:

$$\text{Minimize } \left\{ Z = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i d_{ij} x_{ij} \right\} \qquad (4)$$

Subject to:

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad i = 1, \ldots \ldots \ldots, m \qquad (5)$$

$$x_{ij} \leq x_{jj} \qquad i = 1, \dots \dots \dots, m; \qquad j = 1, \dots \dots \dots, n$$

$$\sum_{j=1}^{n} x_{jj} = p \tag{6}$$

$$x_{ij} = (0, 1) \qquad i = 1, \dots \dots \dots, m; \qquad j = 1, \dots \dots \dots, n$$

Inputs:

x_{ij} = 1 if demand area *i* is assigned to the facility at *j* , 0 otherwise,

i = Index of demand nodes,

m = Total number of demand points in the space of interest,

j = Index of potential facility sites,

n = Total number of potential facility locations,

a_i = Demand at node *i* ,

d_{ij} = Distance between demand node *i* and potential facility site *j* ,

p = The number of facilities to be located,

Equation (4) minimizes the total demand-weighted distance between population and facilities. Equation (5) ensures that every demand is assigned to a facility site i.e. its nearest. Equation (6) limits the number of facilities, p, to be located.

3.1.3.Maximizing Capacitated Coverage

This model is a modification of the Maximize Coverage problem, with the addition of capacity constraints on facilities. Demand is now assigned to the nearest facility with free capacity. Thus, if the nearest facility has reached its capacity limit, demand is then assigned to the second nearest and so on.

This model is primarily concerned with locating centres so that the maximum amount of demand is covered by facilities as long as the capacity limits are not exceeded. Unlike the Maximum Coverage model, it is not compulsory to specify impedance cut offs in running particular problems. If impedance cut offs are set to zero, in ArcGIS 10.1 this problem is then treated as a p-median with capacity constraints. However, if impedance cut offs are specified, the model is equivalent to the Maximum Coverage problem with capacity constraints. This model handles demand as follows:

i. facilities are located so that as much demand as possible is covered within the specified distance, as long as the relevant capacities are not exceeded;
ii. demand points have all or none of their demand assigned to the nearest facility with capacity available;
iii. if the total demand within the impedance cut off is greater than the capacity the facility can receive, only the demand points that would maximize the total captured demand or only those that would minimize the total weighted impedance, when the problem becomes a p-median with capacity constraints, are allocated.

3.1.4. Minimizing Facilities

This model finds the minimum number of facilities required to provide a certain standard of service for the whole region (i.e. within the specified impedance cut off). Effectively, this problem involves running the Maximum Covering problem repeatedly, increasing the number of facilities at each run until all demand is within the covering radius (e.g. 800 m for our schools).

3.2. Data

A full list of the datasets used with detailed descriptions of the procedures used on each is given in the Technical Report. A general outline of the procedures used to estimate the location of pupils, using census data for neighbourhoods and detailed cadastral data, was given in the introduction. Estimating the residential locations of pupils was, in fact, a major problem which consumed much time and effort, mainly because finely detailed spatial data on population could not be obtained from DOSM within the time needed.

3.2.1.The Schools and Their Capacities

The current total number of pupils attending each school with its number of classrooms and the estimated or theoretical maximum capacity of the school, assuming each classroom can accommodate 30 pupils are displayed in Table 1.

Table 1 Current Number of Pupils, the Number of Classrooms and the Maximum Capacity of the School

School	No. of Pupils	No. of Classroom	Max Capacity
SK Rawang	432	16	480
SK Sinaran Budi	929	36	1080
SK Bukit Rawang Jaya	947	30	1800
SK Bandar Baru Rawang	573	30	900
SK Taman Tun Teja	1776	47	2820
Total	4657		7080

3.2.2.Population of Rawang Town by Age and Ethnicity

The distribution of demand was initially estimated by using population data for 23 neighbourhoods in Rawang derived from the 2010 census data collected by DOSM. The primary building blocks of the Malaysian census are Enumeration Blocks (EBs) which generally consist of 80-120 houses or 500-600 people. Unfortunately, after several attempts to obtain population data for EBs in Rawang from DOSM had failed, data which had been provided earlier by DOSM for generally larger areal units defined by the author (termed here 'neighbourhoods') had to be used as the starting point for estimating where pupils live. Though these neighbourhood areas varied in size, they averaged roughly 1670 people and were therefore generally much larger than EBs.

These data from DOSM provide breakdowns of population by gender, age and ethnicity for the neighbourhood areas. Only the three main ethnic groups will be considered here i.e. Malays, Indians and Chinese (Figure 2) since these three are the only ones mentioned in the Malaysia Education Blueprint. The number of pupils from each neighbourhood likely to attend a school was estimated by using data on the relevant age groups and the proportions of each ethnic group attending schools given officially for the country as the whole (respectively 90%, 44% and 8%) (Ministry of Education (MOE), 2012). This number was then distributed evenly across the small cadastral plots or lots within the areas of the neighbourhood classified as residential. These lots or Small Residential Cadastral Plot Units (SRCPUs) can therefore be taken as defining each house or block of flats giving a total of 10,718 demand units or nodes (Figures 3 and 4). The procedure is explained more fully in Section 3.2 of the Technical Report.

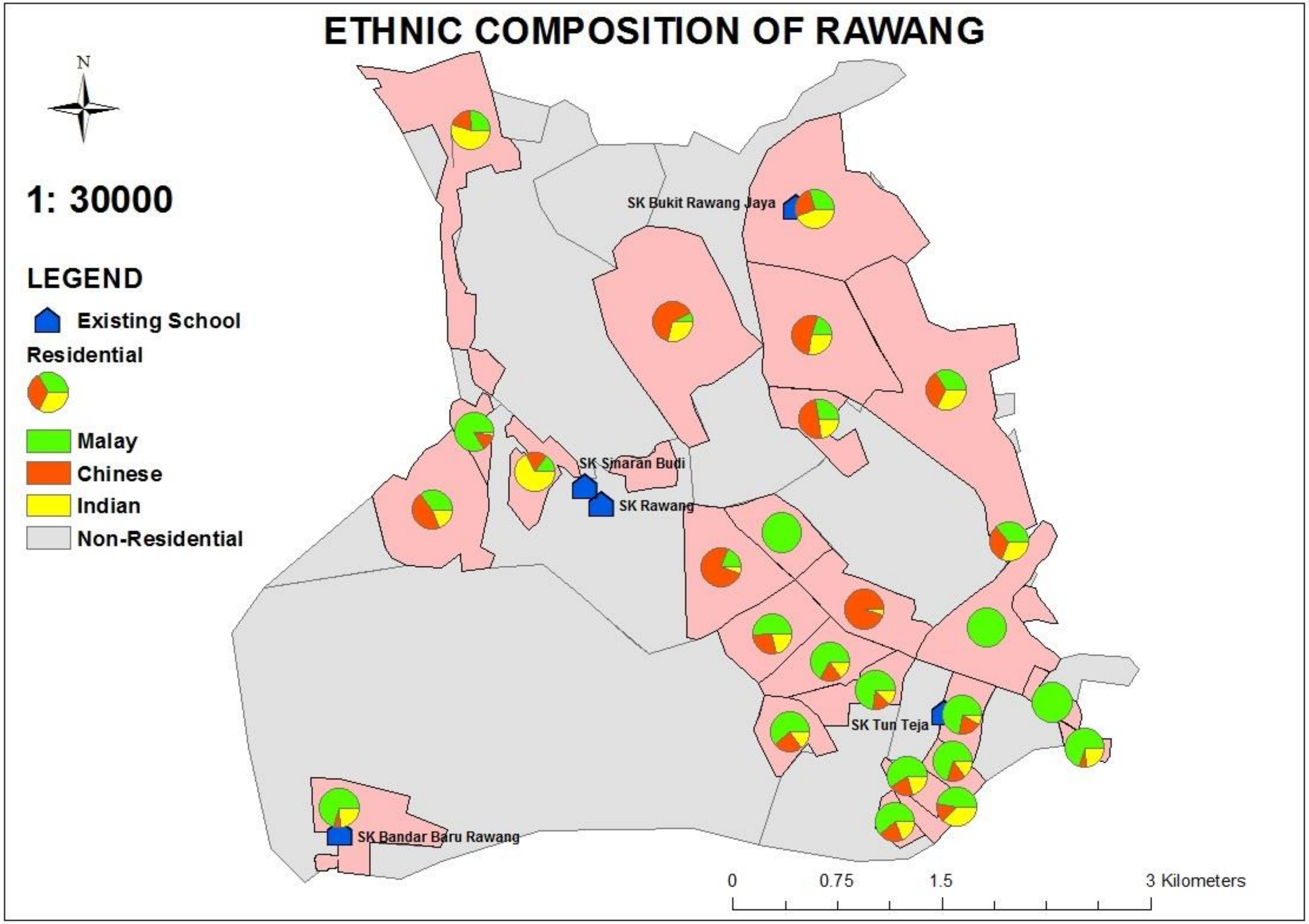

Figure 2 Map of Ethnic Composition of Neighbourhoods in Rawang

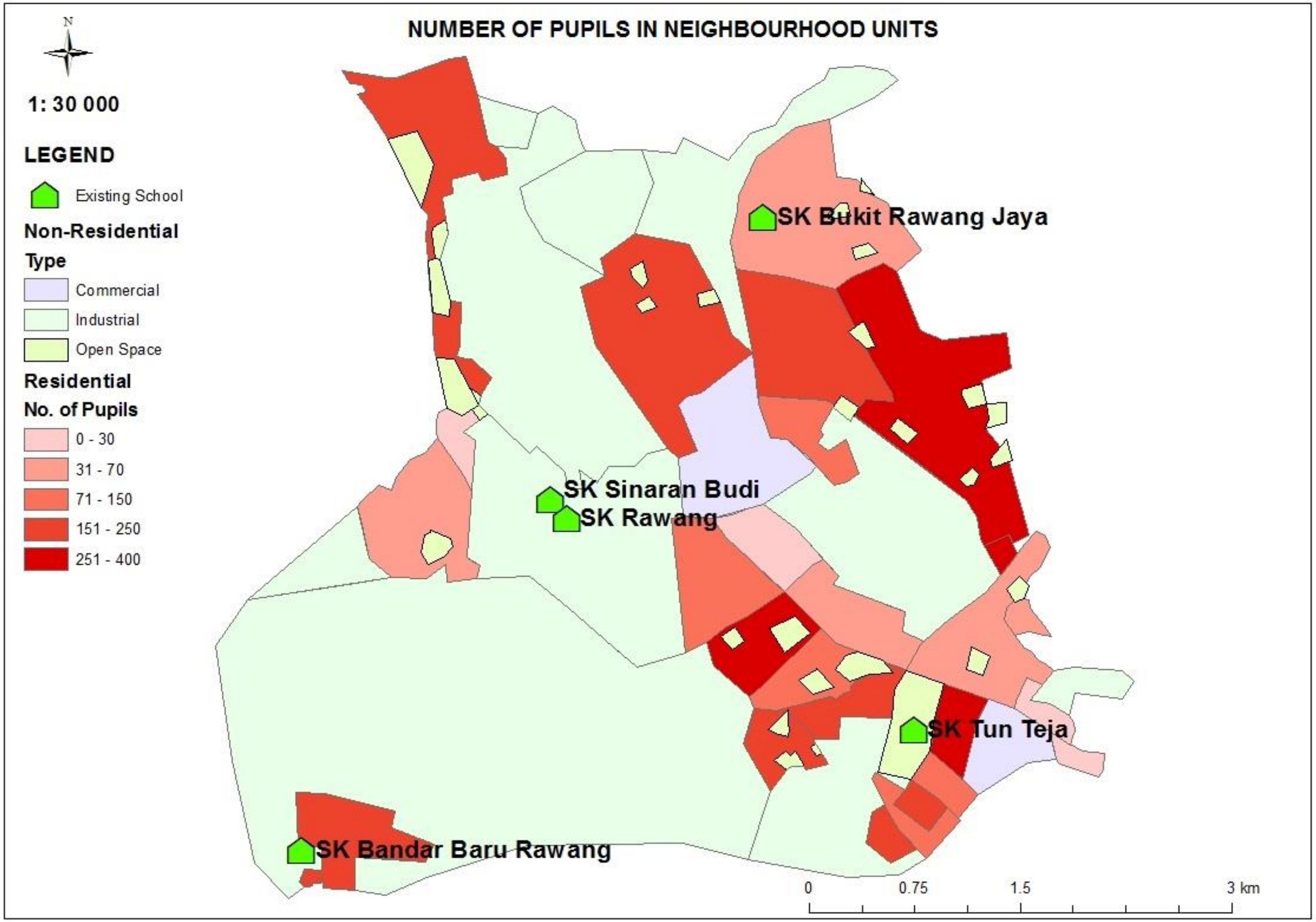

Figure 3 Estimated Number of Pupils in Neighbourhoods

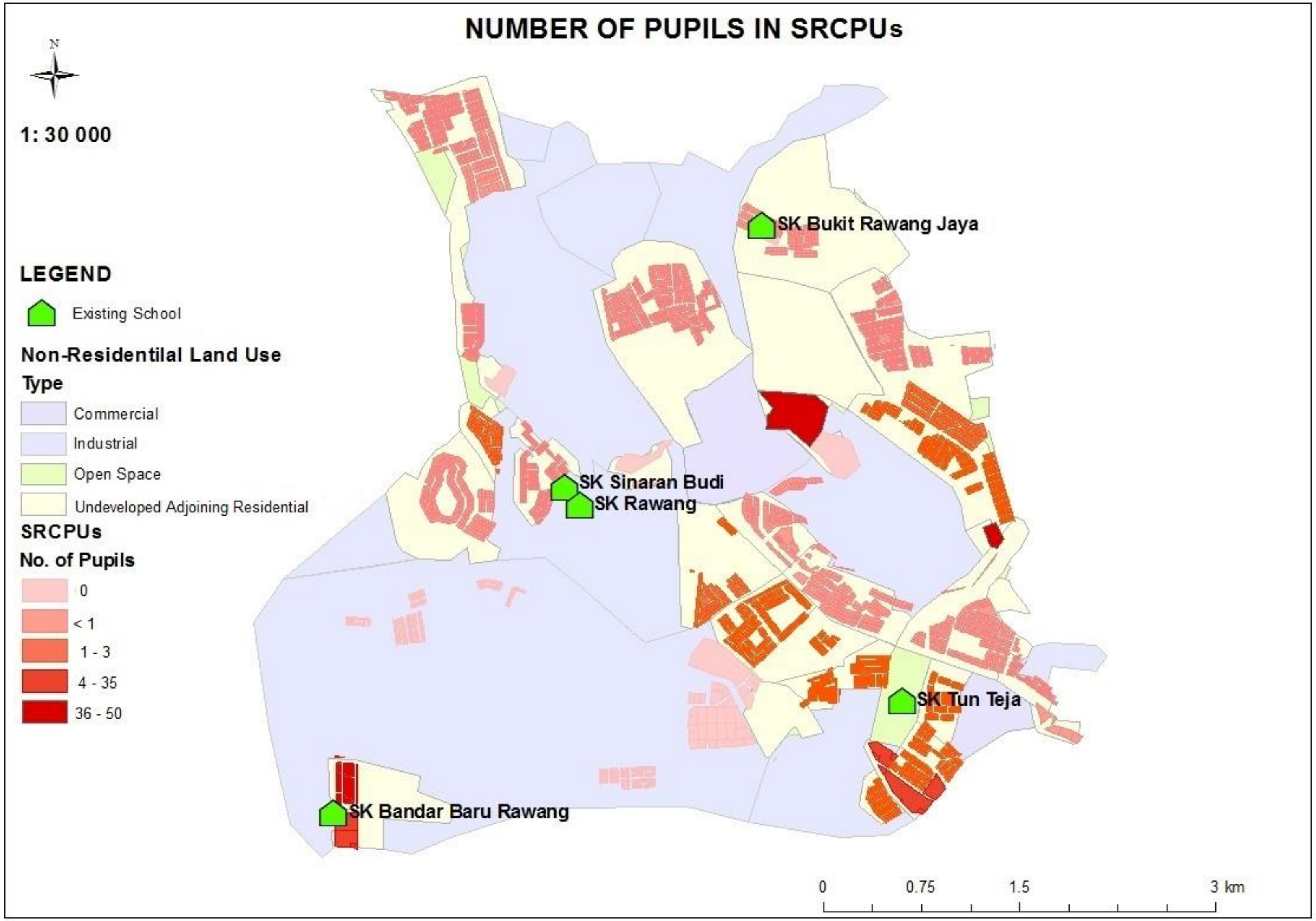

Figure 4 Number of Pupils in Small Residential Cadastral Plot Units (SRCPUs)

4. RESULTS AND ANALYSIS

4.1. Locational Scenarios Examined and Models Used

To assess the existing provision of schools and how it could be improved, four basic scenarios were analysed:

(a) the five existing schools at their present locations;

(b) a selection of the four best located schools from the existing five;

(c) finding the best location for 1 new school to add to the existing 5 ('5+1') and then the best location for adding one more to the latter solution ('5+1+1');

(d) starting with the best four from (b), finding the best location for 1 additional school ('4+1') and then adding 1 further new school ('4+1+1').

When additional schools were being located these were chosen from a set of 27 additional locations located near residential areas where there seemed to be enough open or vacant space to site a primary school. For each scenario 3 LAMs were first run, all without capacity constraints: minimising aggregate (and therefore mean) travel distance; maximizing the number of pupils within 800 m cover distance; and maximizing the number within 1600 m. Then each of these 3 models was run again, but with the appropriate capacity constraints for each school. Thus for each of the 4 scenarios, at least 6 models had to be run.

Despite these being relatively large problems (compared to many LAMs reported with the literature which typically deal with a few hundred demand points at most) as regards the number of demand points (10,718 SRCPUs), the models with no capacity constraints usually only took 5-15 minutes time to run on the computer. However, a few models with capacity constraints sometimes needed 6-7 hours to find solutions if capacity constraints were biting, probably partly because the algorithms used to solve capacitated problems are more complex. Possibly for the latter reason, the results from models with capacity constraints can at first sight sometimes appear puzzling, especially for aggregate or mean travel and therefore have to be viewed with caution. New schools are assumed to have a capacity of 1410 pupils, which would be the capacity of the newest school, SK Tun Teja, if it operated with morning sessions only.

In fact, the results for models with and without capacity constraints were nearly always the same or very similar because, with an estimated 3, 674 pupils in the system but total capacity for 7,080, the capacity constraints generally had little or no effect on the solution process, though it is possible that they may have had some effect during some iterations of the heuristic algorithms e.g. if a particular configuration of schools meant that, say, SK Rawang did not have sufficient capacity for all pupils assigned to it as their nearest school.

4.2. Assessment of the Locations of the Five Existing Schools (Scenario (a))

Figure 5 shows the catchment areas of the existing 5 schools when all pupils are assigned to their nearest school by drawing straight line from each SRCPU to its allocated school. These lines do not trace the actual paths taken through the network; the assignments are determined by real network distances, which explains why some pupils from areas in the extreme north west are allocated to SK Bandar Baru Rawang, presumably via a convenient road, when 3 other schools appear much closer as the crow flies.

Table 2 gives results for the number of pupils covered by each of the existing 5 schools within 800 and 1600 meters. It also gives the number of pupils in each school's catchment when all pupils are assigned to their nearest school, plus the mean travel distance in each catchment. These results seem to indicate that the existing schools are not well located to serve demand. Only slightly over half of the children in Rawang are within 1600 m (1844 or 50.2%) and less than one quarter (902 or 24.6%) are within the recommended distance of 800 m. Thus, school locations seem to be resulting in pupils needing to travel rather far. The mean distance travelled by the children to the five schools is approximately 2.33 kilometres, considerably more than the distance recommended by DTRPS.

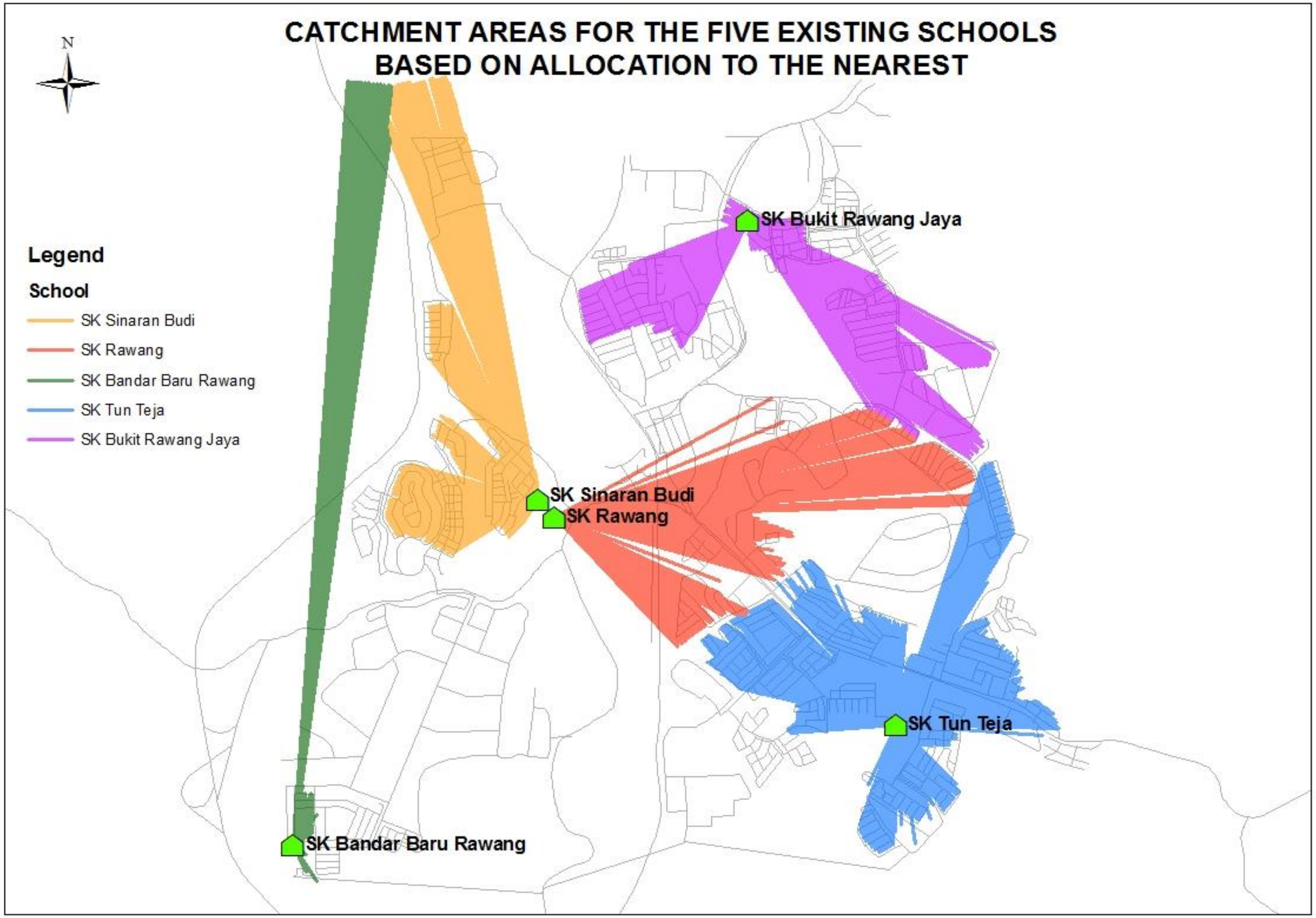

Figure 5 Catchment Areas for Existing 5 Schools

Table 2 Mean Travel Distance to the Existing Schools, Number of Pupils Allocated to Each and Number within the Recommended Distances (Scenario (a))

School	Maximum Capacity	Mean Travel Distance (km)		Number of Pupils					
		With Capacity Constraints	Without Capacity Constraints	With Capacity Constraints			Without Capacity Constraints		
				Within Catchment Area	800 m	1600 m	Within Catchment Area	800 m	1600 m
SK Rawang	480	2.940	2.940	406	0	28	406	0	28
SK Sinaran Budi	1080	1.835	1.835	295	75	120	295	75	120
SK Bukit Rawang Jaya	1800	1.988	1.988	539	35	80	539	35	80
SK Bandar Baru Rawang	900	1.263	1.263	204	173	176	204	173	176
SK Tun Teja	2820	1.568	1.568	2230	618	1440	2230	618	1440
Total	**7080**	**2.233**	**2.233**	**3674**	**902**	**1844**	**3674**	**902**	**1844**

Judging by the pupils covered within 800 m and 1600 m, one school appears to be rather badly located, i.e. SK Rawang which is located just beside SK Sinaran Budi and only covers 28 pupils at the larger distance of 1600 m (only 0.76% of the whole demand). This is misleading partly because when demand points are within a specified covering distance of two or more schools, the LAM used credits only the nearer school for covering these pupils. Since the residential areas near SK Rawang and SK Sinaran Budi are nearly all closer to SK Sinaran Budi, areas which both cover will nearly all be credited to the latter. Additionally, the demand of the adjoining areas is not that high (0.3 pupils at the most per SRCPU) probably because the major ethnic group living nearby is Indian (61%).

SK Bandar Baru Rawang on the other hand is located rather far from most or Rawang's residential areas. Its relatively low average travel distance of 1.263 kilometres suggests it is generally serving pupils from areas nearby, as Figure 5 suggests. Beyond that there are extensive zones of industrial land use. This explains its small catchment of 204 and why the number of pupils it covers only increases from 173 to 176 when the cut off rises from 800 m to 1600 m.

SK Tun Teja has far more pupils within 800 m and 1600 m (618 and 1440) in both sets of results, so it is providing relatively good access to substantial numbers of pupils in residential areas fairly close (50-70% Malay and having a number of high rise apartment blocks), confirmed by the mean catchment distance of 1.568 km. While the equivalent results for SK Bukit Rawang Jaya give it only 35 and 80 pupils within these cover distances, from Figure 5 it appears to be serving a sizable catchment spread across the east and north east of Rawang (539 pupils with a mean distance of 1.988 km) but much less than its capacity of 1800 could accommodate.

SK Rawang does have a sizable catchment (406), though it is much less compact the others (Figure 5), with a mean travel distance of 2.940 km and only 28 pupils within 1600 m. SK Sinaran Budi has the second smallest catchment area of the 5 (295) and only has 75 and 120 pupils within 800 m and 1600 m respectively. These less favourable results mainly reflect its sharing a fairly central location with SK Rawang.

The notional coverage areas of these 5 schools and a version of the land use map which highlights more dense residential areas are shown together in Figure 6. Three broad areas with high demand and poor accessibility (red in Figure 6) stand out outside the 1600 meters coverage of the existing schools. Most of the 27 candidate locations mentioned earlier are located in or near these three areas. Numbers identifying these 27 candidates are also given in Figure 6.

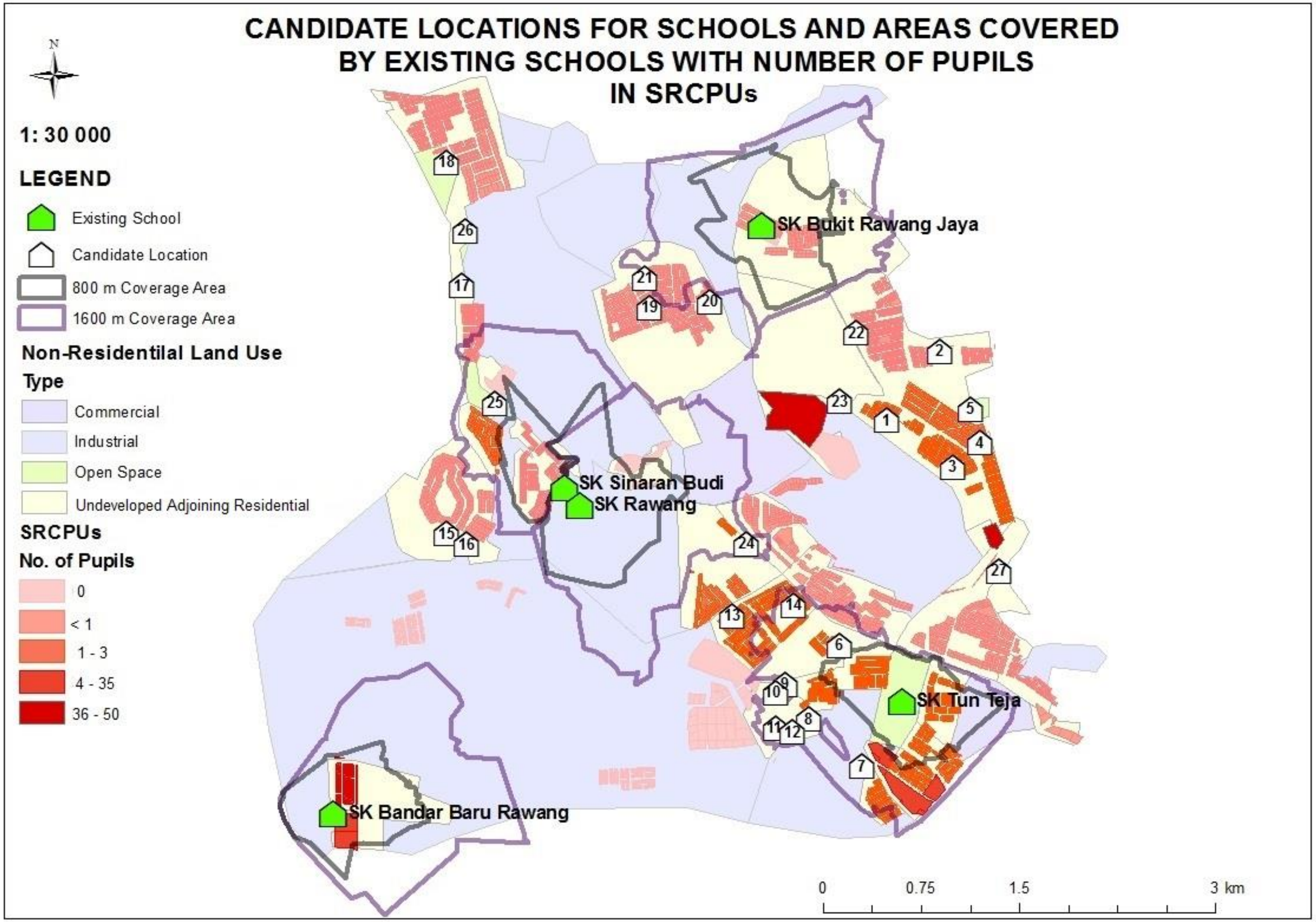

Figure 6 Candidate Locations for Additional Schools and Areas Covered by Existing Schools with Number of Pupils in Small Residential Cadastral Plot Units (SRCPUs)

4.3. Selecting Four out of the Existing Five Schools (Scenario (b))

Scenario (b) is designed to explore the possible consequences of closing the least well located of the 5 existing schools (Figure 7 and Table 3) with the existing 5 treated as candidate sites in selecting the best 4. The school not selected is then the one that could be considered for closure.

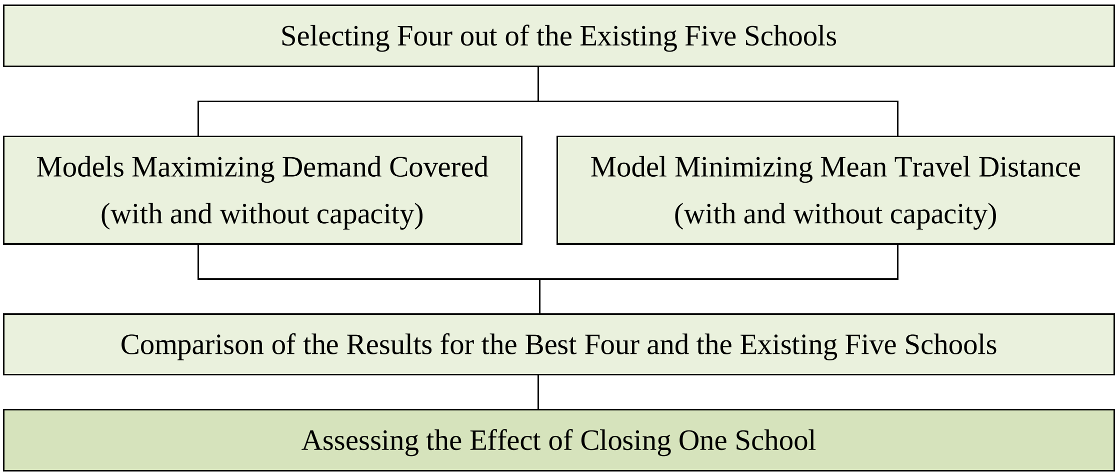

Figure 7 LAMs to Explore the Possibility of Closing One School

Table 3 Comparison of Aggregate Travel Distance and Number of Pupils Covered for Scenarios (a) and (b)

Number of Schools	**With Capacity Constraint**			**Without Capacity Constraint**		
	Aggregate Travel Distance (km)	**No. of Pupils Covered**		**Aggregate Travel Distance (km)**	**No. of Pupils Covered**	
		800 m	**1600 m**		**800 m**	**1600 m**
Existing Five	22, 735.596	902	1844	22, 735.596	902	1844
Best Four	23, 012.652	902	1817	23, 336.941	902	1834
Differences	**277.056**	**0**	**-27**	**601.345**	**0**	**-10**
School Selected to be Closed	**SK Rawang**	**SK Rawang**	**SK Rawang**	**SK Sinaran Budi**	**SK Rawang**	**SK Sinaran Budi**

Not surprisingly, SK Rawang is selected for 4 of the 6 models as the school to be closed with SK Sinaran Budi selected for the other two. SK Rawang has the lowest numbers of pupils covered at 800 m and 1600 m and the largest mean distance for its catchment, though it fares better in size of catchment. While SK Sinaran Budi fares better on all of these criteria, except catchment size, it has the 3rd worst covering values and the second smallest catchment. Since it is very close to SK Rawang, it therefore seems inevitable that it should be selected in the two models where SK Rawang is not selected.

The reason why SK Rawang is closed in the 3 models with capacity constraints is because it has a smaller capacity (480) than its neighbour (1080). Thus the latter can serve all the demand of the surrounding areas accessibly if SK Rawang is closed. However, the reverse is not the case and the loss of SK Sinaran Budi's capacity would force some pupils to travel to schools further afield with longer distances.

As Table 3 shows, there is not much difference to the aggregate distance in the whole system when the number of schools is reduced to four in one of the preceding ways, because the two prime candidates for closure are located just beside each other. That is also why aggregate travel distance in the uncapacitated model does not increase much. Similarly, there are only very slight reductions in the number of pupils covered within 1600 m by the four selected as opposed to the five existing schools. If some re-organisation is being considered with one of the schools near the centre being closed and provision expanded in the outer areas, these results suggest that SK Rawang is the stronger candidate for closure.

4.4. Location of Additional Schools

Scenarios (c) and (d) are designed to explore the benefits of adding 1 or 2 schools to the existing provision (Figure 8). Scenario (d) uses the best 4 of the existing schools as its starting point with SK Rawang always selected for closure. In both scenarios (c) and (d), the best school from the 27 candidates to add to the existing 5 or best 4 is selected first. Then, the next best additional school is selected from the 26 candidates left to add to the 6 or 5 now in the solution.

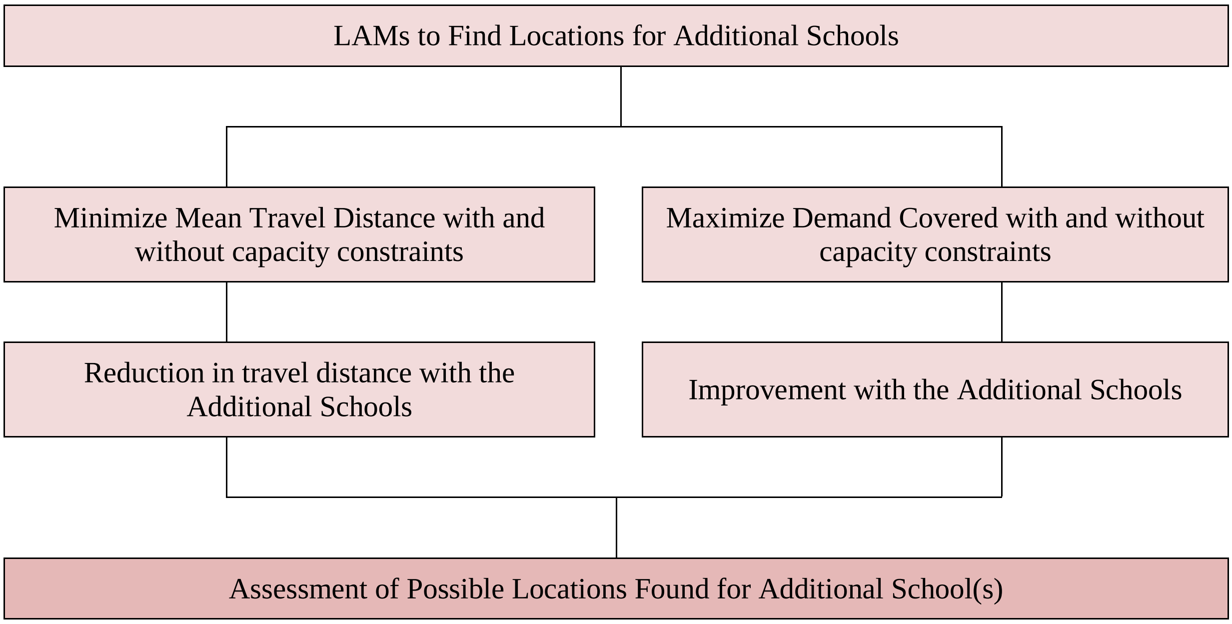

Figure 8 LAMs to Find Locations for Additional School(s)

4.4.1. Adding 2 Schools to the Existing Five (Scenario (c))

Table 4 sets out the main results for scenario (c). There are substantial improvements in all the criteria used when one school is added and again when a further school is provided, with the latter improvements generally ranging from only slightly smaller for pupils within 800 m to significantly smaller for the other covering objective, both for capacitated and uncapacitated problems (Table 4). Even when 2 schools are added, only 1,715 pupils (46.7%) are within 800 m of a school, though 2,983 (81.2%) are within 1600 m for both capacitated and uncapacitated problems. An exception to this pattern of improvements is the capacitated p-median problem where the second additional school (18) brings a much larger reduction in overall travel distance than the first (4).

Table 4 Improvement in Location-Allocation Results with One or Two Schools Added to the Existing Five (Scenario (c))

LAMs			Existing 5	Add 1	Improvement	New School Chosen	Add 1 More	Improvement	New School Chosen
With Capacity Constraints	Aggregate Travel Distance (km)		22, 735.596	21, 045.835	-1, 689.761	4	16, 249.417	-6, 486.179	18
	Number of Pupils Covered	800 m	902	1, 344	442	27	1, 715	813	14
		1600 m	1, 844	2, 671	827	4	2, 983	1, 139	24
Without Capacity Constraints	Aggregate Travel Distance (km)		22, 735.596	18, 575.082	-4, 160.514	4	16, 871.533	-5, 684.063	14
	Number of Pupils Covered	800 m	902	1, 344	442	27	1, 715	813	14
		1600 m	1, 844	2, 671	827	4	2, 983	1, 139	24

Table 5 Review of Schools Selected as Additional to the Existing Five

Selected Additional School	**Rank when 2 Additional Schools are Selected**						**Demand of Surrounding Area (Per Housing Unit)**					**Remarks**
	With Capacity Constraint			Without Capacity Constraint								
	Minimizing Distance	800 m	1600 m	Minimizing Distance	800 m	1600 m	0	< 1	1 – 3	4 - 35	36 - 50	
4	1st		1st	1st		1st		✓	✓		✓	Surrounded by residential areas with high rise apartments nearby
14		2nd		2nd	2nd			✓	✓			- Between SK Tun Teja and SK Rawang catchment areas - High proportion of Malays in several areas nearby
18	2nd							✓				Located far from other existing schools making it cover an area where pupils are poorly served.
24			2nd			2nd		✓	✓			- Near to a residential areas which is 100% Malay - Close to SK Rawang's catchment area
27		1st			1st			✓		✓		Located far from other existing schools making it cover pupils that are poorly served and near to a residential areas which is 100% Malay

In the first stage candidate 4, positioned in the Rawang Perdana 2 neighbourhood near the middle of the poorly served eastern part of Rawang, is selected for both capacitated and uncapacitated versions of the p-median problem and also for both versions of the second covering problem. Candidate 27, in the southern part of the latter poorly served eastern area is selected to maximise pupils covered within 800 m in both versions of the latter problem, presumably because it has more pupils in its immediate vicinity than 4.

In the second stage of the latter problem candidate 27 is paired with 14, located roughly in the centre of the poorly served area between SK Tun Teja and SK Rawang, for both versions of this problem. To maximise pupils covered within 1600 m, in both versions of the problem candidate 4 is paired with 24 about half a kilometre north west of 14. In the capacitated p-median problem with 2 extra schools, candidate 4 is paired with location 18, located in the poorly served neighbourhood of Taman Garing Permai in the extreme north west (Figure 9), whereas for the uncapacitated version is paired with 14, also in a somewhat poorly served area, as just noted. Thus, both candidates chosen for the 6^{th} school and the various candidates they are paired with in providing the 7^{th} seem to fill significant gaps in the existing pattern of spatial provision, which gives visual confirmation that the LAMs are choosing good locations.

4.4.2. Adding Two Schools to the Best Four of the Existing Five (Scenario (d))

The results for (d) are described in Table 6. Since SK Rawang is the school closed in scenario (d), the locational configurations of (d) and (c) are very similar. Consequently, nearly all the results in Table 6 regarding the new sites selected at each stage and the improvements they produce in both versions of the 3 models are very similar, or even identical. The only notable difference is that the addition of candidate 4 in the first stage of the capacitated p-median problem reduces aggregate travel distance much more than it does in scenario (c), though the combination of candidate 4 with 18 in the second stage brings a similar improvement to that achieved by this combination in scenario (c).

The one contrast in these sets of results here appears somewhat puzzling initially, but is probably mainly due to the differences in the capacities available to central areas of Rawang when only SK Sinaran Budi is open in (d), as opposed to both the latter and SK Rawang being open in (c).

With one exception, the same pairs of candidates are selected as before by the various models so that they complement each other by locating in poorly served areas some distance from each and thus filling different gaps in supply. The exception is candidate 7 which replaces 14 as the second location for the first capacitated cover problem, though in fact it is quite close to 14 (less than 700 meters from it).

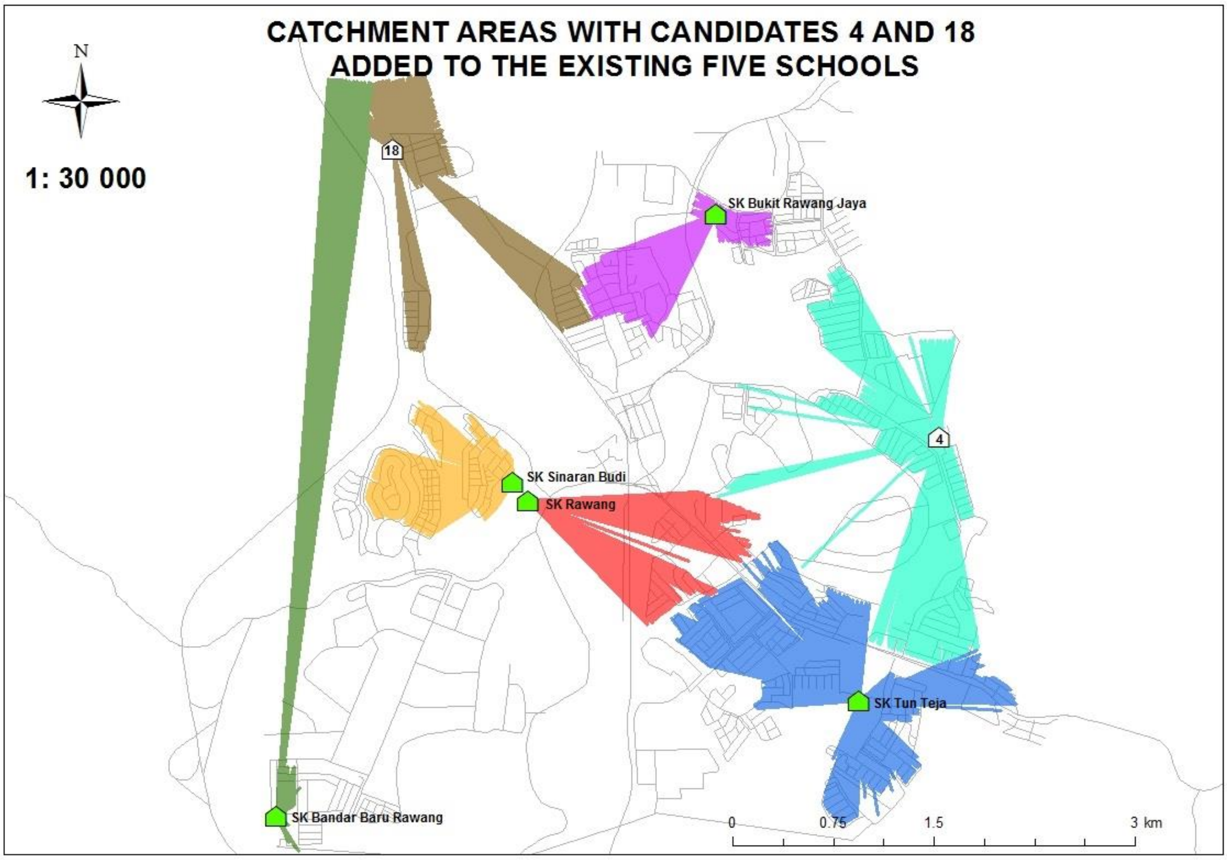

Figure 9 Catchment Areas with Candidates 4 and 18 Added to the Existing Five Schools

Table 6 Improvement in Location-Allocation Results with One or Two Schools Added to the Best Four of the Existing Five (Scenario (d))

LAMs			Existing 4	Add 1	Improvement	New School Chosen	Add 1 More	Improvement	New School Chosen
With Capacity Constraint	Aggregate Travel Distance (km)		23, 012.652	18, 993.860	**-4, 018.79**	4	16, 413.002	**-6, 599.65**	18
	Number of Pupils Covered	800 m	902	1, 344	**442**	27	1, 713	**811**	7
		1600 m	1, 817	2, 643	**826**	4	2, 983	**1, 166**	24
Without Capacity Constraint	Aggregate Travel Distance (km)		23, 012.653	18, 739.043	**-4, 273.61**	4	16, 873.671	**-6, 138.982**	14
	Number of Pupils Covered	800 m	902	1, 344	**442**	27	1, 715	**813**	14
		1600 m	1, 817	2, 643	**826**	4	2, 983	**1, 166**	24

5. CONCLUSION

In our analysis, the location of demand and the journey to school were both treated in a more detailed and realistic manner spatially than has been the case in virtually all equivalent problems found hitherto in the recent research literature, a significant contribution of the present work and a benefit of having a wide range of LAMs integrated into a GIS with network capability. At this fine scale with 10,718 demand nodes, DTRPS's recommendation that all children should be within 800 m of a school seems over-ambitious as only 24.6% of the estimated pupils attending national schools in Rawang enjoy this level of access to the existing 5 schools. Even with 2 additional schools added at good locations, this only increases to 46.7% in scenario (c). A distance of 1600 m would serve as a more achievable goal: 50.2% of our pupils are within that distance of the 5 existing schools and 81.2% would be with 2 more well located schools in scenario (c). Results from the Minimizing Facilities model suggest, however, that to cover all pupils within 1600 m would require 12 schools, which is not feasible practically.

The LAMs applied produced results on a range of criteria which gave insight into how well the 5 existing schools serve the area, as well as where substantial areas of poorly served demand lie and also helped to identify and evaluate some good locations for siting one or two new schools. Several pairs of candidates seemed to merit particular consideration and further evaluation in this context, namely candidates 4 and 18 or 4 and 14 or 4 and 24. Candidates 4 and 18 also have the advantage of being located near new residential areas, largely developed since 2010, and therefore not included in our demand data.

Being able to run models with capacity constraints was generally helpful as it could be illuminating to compare results with and without such constraints. Though the two versions of the models rarely produced significantly different results here, they did produce different outcomes in scenario (b) for reasons explained earlier. Generally, scenario (c) seems more realistic than (d), since SK Rawang is unlikely to close because it is the oldest school in Rawang, is popular with parents and alumni and enjoys a good reputation, as well as a central location.

The results from the LAMs used here could be further enhanced if demand could be estimated more accurately using data on EBs or even smaller user defined units plus data on new residential areas, as well as information on how many children from Rawang go to schools elsewhere (e.g. in Kuala Lumpur or rural areas where their grandparents live) and on how many children travel to schools in Rawang from outside. Further ahead, more sophisticated models could be developed taking account of a wider range of variables such as actual modes of travel to school and preferences for different schools.

REFERENCES

Abdul Rahman, B. (2014, January 13). How Students are Allocated to the Schools? (N. Naharudin, Interviewer)

Ariffin, M. (2011). *Location Allocation Problem Using Genetic Algorithm and Simulated Annealing: A Case Study Based on Schools in Enschede [Master Thesis]*. Faculty of Geo-Information Science and Earth Observation: University of Twente.

Blazevic, A. (2006). *Using Location-Allocation to Optimize the Location of Fire Tanker Airbases in California: Master Thesis.* Department of Geography and Environmental Resources: Southern Illinois University Carbondale.

Chang, K. T. (2012). *Introduction to Geographic Information Systems.* New York: McGraw-Hill.

Department of Statistics Malaysia (DOSM). (2010). *Population Census Data for Rawang.* Putrajaya: Department of Statistics Malaysia (DOSM).

Department of Town And Regional Planning Selangor. (1999). *Rawang-Kuang Local Plan* (2nd ed.). Shah Alam: Department of Town And Regional Planning, Selangor.

Department of Town and Regional Planning, Selangor. (2010). *Garis Panduan & Perancangan Piawaian Perancangan Negeri Selangor [Guidelines & Planning Policies for the State of Selangor]* (2nd ed.). Shah Alam: Department of Town and Regional Planning, Selangor, Malaysia.

ESRI. (2010). *Approaches to School of Business GIS Programs*. Retrieved 2014 from ESRI: http://www.esri.com/library/brochures/pdfs/approaches-to-school.pdf

ESRI. (2012). *Network Analyst.* Retrieved 2014 from ESRI: http://www.esri.com/software/arcgis/extensions/networkanalyst/key-features

Goodchild, M. (1992). Geographical Information Science. *Geographical Information Systems, 6*(1), 31-

Hakimi, S. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research, 13*, 462-75.

Hakimi, S. (1965). Optimum Distribution of Switching Centres in a Communications Network and Some Related Graph-Theoretic Problems. *Operations Research, 13*(462-75).

Handayani, L. (1997). *The Application of Geographical Information Systems (GIS) to School Planning Problems: Selected Examples.* Thesis MAppSc: University of New South Wales.

Hodgart, R. L. (1978). Optimizing Access to Public Services: A Review of Problems, Models and Methods of Locating Central Facilities. *Progress in Human Geography March, 2*, 17-48.

Hodgart, R. L. (1981). *On the optimal organization of public services with central facilities: PhD Thesis.* University of Edinburgh.

London, N. A. (2007). School Location in a Developing Nation: A Location-allocation Modelling Approach. *International Journal of Educational Management, 4*(5).

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2011). *Geographic Information Systems & Science* (3rd ed.). Danvers: John Wiley & Sons. Inc.

Longley, P., & Batty, M. (2003). *Advanced Spatial Analysis: the Case Book of GIS.* California: ESRI.

Lwin, K. K., & Murayama, Y. (2012). GIS Network in Geospatial Analysis. In *Progress in Geospatial Analysis* (p. 183). Japan: Springer.

Marianov, V., & Serra, D. (1998). Probabilistic, Maximal Covering Location-Allocation Models for Congested Systems. *Journal of Regional Science, 38*(3), 401-424.

Marianov, V., & Serra, D. (2002). Location Problems in the Public Sector. In *Facility Location Application and Theory* (pp. 119-143). New York: Drezner, Z. & Hamacher, H.

Ministry of Education. (2012). *Malaysia Education Blueprint.* Putrajaya: Ministry of Education, Malaysia.

Møller-Jensen, L. (1998). Assessing Spatial Aspects of School Location-Allocation in Copenhagen. *Geografisk Tidsskrift-Danish Journal of Geography*, *98*(1), 71-80.

Pearce, J. (2000). Techniques for Defining School Catchment Areas for Comparison with Census Data. *Computers, Environment and Urban Systems*, *24*, 283-303.

Schietzelt, T., & Densham, P. (2003). *Location-allocation in GIS.* GIS Applications.

Turner, D. (2006). *Implementing the Flow-Interception Location Model with Geographic Information Systems.* Master Thesis: University of Texas at Dallas, USA.

Yeh, A., & Chow, M. (1996). An Integrated GIS and Location-Allocation Approach to Public Facilities Planning- An Example of Open Space Planning. *Elsevier Science, 20*(4-5), 339-350.

PART II

TECHNICAL REPORT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DOSM	Department of Statistics, Malaysia
DTRP	Department of Town and Regional Planning
GIS	Geographical Information System
JUPEM	Department of Surveying and Mapping
LAM	Location-Allocation Model
LAP	Location-Allocation Problem
MOE	Ministry of Education
NDCDB	National Digital Cadastral Database
SRCPU	Small Residential Cadastral Plot Unit
TIN	Triangular Irregular Network

1. FURTHER REVIEW OF LITERATURE AND OF METHODS

1.1. Location-Allocation Models

The decision whether to locate, relocate or increase the number of public facilities to support ta growing demand is always a challenge to the government in most of the developing countries, including Malaysia. Location-Allocation Models (LAMs) can be used as a support for the decision making process in improving the spatial location of public facilities to enhance the spatial equity or efficiency of the service in meeting the demands of users (Buzai, 2013).

Finding optimal locations for facilities can include various graph-theory techniques (e.g. finding shortest paths), methods for solving the p-median problem or suitability mapping etc. (Hakimi, 1964). The integration of p-median problems with land use maps, for example, could provide a set of suitable candidate locations for facilities at promising locations for serving demand efficiently. The p-median problem could also be combined with advanced spatial models that include various criteria such as land use classification or land cover type and features of the intended facilities in locating facilities to provide good accessibility for users and accommodate demand better (Abdel-Latif, 2007). For example, in Malaysia, new schools will generally be built at places with close proximity to residential areas and with sufficient space for buildings and other amenities like canteens, playing fields and other sports facilities.

It is essential to consider the transportation system and movement through the network which connects different resources and demands throughout an area when employing LAMs. For instance, if there is a limited number of facilities to serve the whole population, LAMs can be used to find good solutions which allocate as much population as possible to facilities which are close or accessible to their users. LAMs can help in solving the problem of matching supply and demand spatially by incorporating various constraints e.g. on how far users can feasibly travel. It is important to try to provide an equitable service to population while maximizing efficiency through minimizing the total distance travelled or time taken. Location-allocation models require information on the location of supply, usually treated as point locations, and on demand - that is the locations of individual users or consumers plus impedance measures which can be distance or time travelled. Distance or time can be measured by the shortest path to supply centres in a network which is, of course, more accurate and realistic than straight-line distances (Chang, 2012).

LAMs are usually set up for a specified number of facilities and can involve using constraints which specify maximum distances or costs of travel and possibly maximum distances from demand points to facilities. After the analysis is completed, a cost-benefit analysis can be done by examining the reduction in distance travelled resulting, for instance, from adding one or two new facilities. In this study, LAMs were initially used to evaluate how well located existing schools are to serve the existing distribution of children. Other models were then run to explore various LAM-related problems, for instance, models which ensured all children in the study area enjoyed a certain desirable or standard of access to school.

1.1.1. Previous Studies on Location-Allocation Models for Schools

In a case study done in Rascht City, Iran, school catchment areas were created based on the spatial distribution of students and the capacities of schools (Ahmadi Nejad Masouleh, 2006). The boundaries of each school's catchment were constructed so that the students were allocated to the facility closest to their home. LAMs can also be used to maximize the attendance at each school and match its capacity and also to reduce the distance travelled by students to school.

Maps of school catchment areas can be very useful in identifying areas that are not served well, such as areas that are located too far from the existing schools (Grip, 2013). Existing catchment areas can be modified to suit the changing character and distribution of demand possibly due to growing population. The distribution of supply centres can be evaluated partly through identifying the pupils who are not allocated to the closest school at present. LAMs can help in redistricting the catchment areas to allocate more pupils to their closest school.

When LAMs incorporate capacity constraints, any overcrowding problems of existing schools can also be addressed as well. This method has been used in Baltimore for planning a new school to address the overcrowding of the existing schools (Crepper GIS, 2013). By determining the schools that exceeded their capacity limits, the new school location was planned taking account of the boundaries between overcrowded schools where the new school could capture the surplus demand from its neighbours. Essentially, LAMs simultaneously locate schools and allocate pupils to the appropriate school.

1.2. The Distribution of Demand

1.2.1. Estimation of Demand

Population is a key component in assessing demand for facilities. It is desirable to have population data at the finest scale possible in order to precisely capture where the demand for the service is located in an area (Zandbergen & Ignizio, 2010). The population census in some countries, such as Malaysia, is aggregated to state or local government areas. The planning of schools seems to be mainly based on the population of these large areas. There is a need, however, for such rather large zones to be disaggregated, creating smaller areal units for population. This should lead to better planning of the location of facilities to serve demand. In this study, the method used as a first step in trying to estimate the location of population at a finer scale is the dasymetric technique. Population figures for smaller areas are estimated through using the boundaries of other types of aggregated information, in this case data on types of land use (Cai, et al., 2006). This dasymetric technique allows the different land use types within a two-dimensional census zone to be represented. The non-residential areas are then masked out to distribute the population only to the residential parts of the zone. It has been used previously in disaggregating population census data to smaller areas in Leicester, UK (Mohammed, et al., 2012). Another technique that could have been used in this context would have been to use the road network to allocate census population totals directly to housing units based on streets or houses or groups of houses. However, in this study, only the dasymetric approach is used, because it allows areas identified as residential on the land use map to be conveniently overlaid on cadastral data for small lots or plots, thereby identifying lots occupied by houses or apartment blocks.

Figure 1 shows the flow diagram for the process of estimating the population at a finer scale by using the dasymetric method, as suggested by Mohammed et al. (2012). The non-residential land use is first masked out from the land use map leaving only residential areas on the map. A map of cadastral plot boundaries derived from a plan of cadastral lots is then combined with the residential land use map to identify which cadastral plot units are likely to be residential. Finally, the census population of each ‘neighbourhood’ is distributed appropriately to each of the cadastral units.

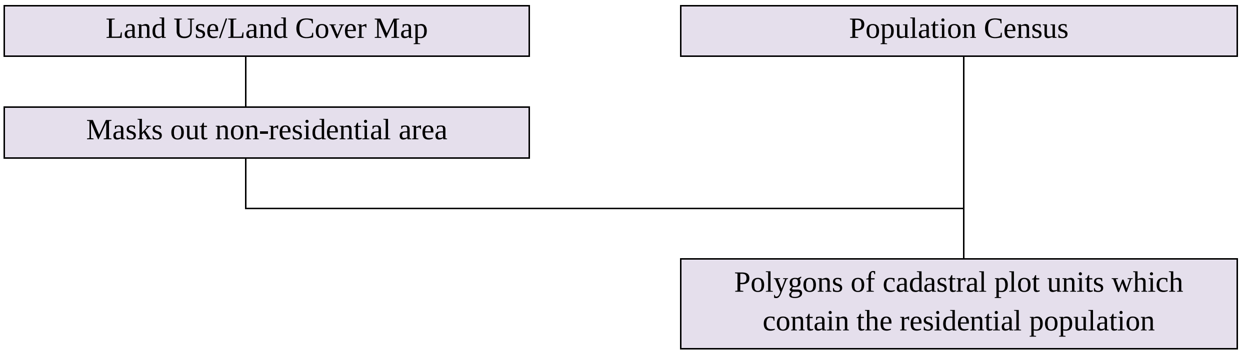

Figure 10 A Dasymetric Technique for Estimating Population Distribution at Finer Scale

1.2.2. Visualization of data by using Choropleth Maps

A Choropleth map is a cartographic-based thematic mapping technique for presenting and visualizing the statistical data of an area. Each area will be shaded according to the statistical value it holds. It is a widely used technique for visualization, for example, population density maps. It can be linked to the dasymetric technique to improve the accuracy of choropleth map, for instance by shading only the area within a zone or unit which is known to be populated and leaving unshaded unpopulated areas like parks or lakes within the zone in question, where the colour can be scaled according to the density value of each region (Andrienko, et al., 2001).

The numerical values need to be classified in order to be clearly differentiated on the map. There are five types of classification methods commonly used which are Natural Breaks (also known as Jenks Natural Breaks), Equal Interval, Quantiles, Standard Deviation and Geometric Progression, each of which classifies the range of statistical values differently, as the names suggest. The most commonly used technique is probably Natural Breaks to show the discontinuities natural where the statistical values are then divided according to the number of classes needed to capture these breaks (Cote, 2012).

Choosing an appropriate colour progression to give a good visual impact on the map is another challenge with choropleth maps. The colour progression needs to communicate any patterns well to the map reader. There are basically six techniques of hue progression in displaying the data on the choropleth map which are single-hue, bipolar, blended hue, partial spectral, full spectral and value progression (Cote, 2012). A good choice of colour progression can enhance effective communication between the map and the reader, but a poor selection can result in an ineffective and unappealing map.

1.3. Network Models

GIS provides various data models which facilitate spatial analysis, including models for road networks, river networks or utility networks. The most important network model for solving problems in urban areas is the road network model. This model is widely used in transportation planning, retail market analysis, service allocation, measuring accessibility and so on. It allows simple daily problems to be solved such as finding a shortest route between two locations, looking for the closest facility etc. (Lwin & Murayama, 2012).

In GIS, a network is a set of nodes and links which are topologically connected to each other. This means a network should allow flow characteristics to take place, such as the movement of people, vehicles or goods in a road network (Chang, 2012). Networks contain a set of nodes, representing the intersections of links in the network, and links which connect two or more nodes. When the tables of nodes and links are relationally connected, this creates a topological relationship between both elements.

Network elements can have weights associated with them, for instance the cost of traversing a link in the network in terms of distance or time or cost can be associated with that link. In addition, in transportation planning, the road network can include a speed limit or other road restriction as a component of cost. This weight can be a decisive factor in finding the shortest or least cost path between two points. Consequently, such data models can be used in various applications such as geocoding and routing problems.

For Rawang, a network model was built from the road network data for the city with the nodes consisting of the intersections and junctions of the network. The weight or cost element is the length of the links between the nodes. In the analysis, left and right turns and u-turns are allowed at the nodes. Subsequently, the road network was topologically connected to the facility locations (schools) and to units of demand i.e. the small areal units of the cadastral plots used to estimate where children live.

1.4. School Catchment Areas

The location of schools should be such as to satisfy the demands of the local school-age population. It is therefore crucial to include census data in any analysis determining the existing school catchment areas across the neighbourhoods of the city. Ideally, this approach should help to describe and assess the present and future distribution of pupils in the neighbourhoods around schools (Pearce, 2000). This can then lead on to an analysis of the school capacities and whether they suit actual or future demand or not.

1.4.1. Thiessen Polygons

Thiessen polygons are such that any point within this type of polygon is closer to one of the set of initial ‘seeds’ (initially specified points or known points) than to the other seeds in the set of seeds. Also called Voronoi polygons, these polygons can be used in various applications, especially for service area analysis for public facilities. By using initial triangulation, the seed points can be connected to form sets of triangles. Delaunay triangulation, for example, ensures that each seed point is connected to its nearest neighbours and the triangles are as equilateral as possible. Once triangulation is done, the Thiessen polygons can be constructed by connecting lines drawn perpendicularly to the sides of the triangles at the midpoints of these sides (Chang, 2012). Thiessen polygons are used internally in GIS to speed up certain geometric operations and as the basis of some powerful methods for generalizing vector databases.

In order to solve the Location Allocation Problem (LAP) for schools, Thiessen Polygons can be used as initial rough indicators in estimating the notional catchment areas of the schools or their ‘natural catchments’. The schools can be regarded as centroid points and the polygons will then be created by connecting each centroid or centre to the residential areas for which it is the nearest centre. Areas not covered by such polygons can then be observed.

This method has been demonstrated by Pearce (2000) where the catchment areas of the schools in Lancashire were determined by using the Voronoi-points generated from Thiessen Polygons. The boundaries of the polygons were calculated by finding the perpendicular bisectors of the lines connecting two neighbouring central points i.e. schools. In this type of analysis, the catchment areas of the schools depend on the spacing between the points. This method is useful here as children in Malaysia tend to go to the school which is located nearest to their home, though this method of allocation is rather simplified compared to real world school catchment areas which are influenced by school capacities, among other factors.

1.4.2.Isochrone Maps

Isochrone maps identify and show the areas around facility locations on a network consisting of all the demand points or street units lying within a specified travel time (or distance or cost) of the facility. Accessibility refers to how easy it is to get to a site and can be measured by distance or time travelled or any other impedance on a network.

A simple way to evaluate accessibility is by setting a buffer defined by Euclidean distance around a point. However, considering that people travel by road, this method does not accurately reflect the actual accessibility of the site. Identifying the streets accessible within certain distances of a site via the road network can overcome this limitation. Once created, a network space allows the catchment areas of the facilities to be identified. Multiple concentric service areas show how accessibility changes in terms of the points, lines or polygon units making up the catchment area with set increases in impedance (Chang, 2012).

In this study, the application of the Isochrone method to identify travel distance contours can help in determining the coverage area of schools. It uses the distance through the road network to identify the area and its population that is within a specified network distance of a school which is an intriusic procedure in the maximum coverage location problem (MCLP).

2. REVIEW OF DATASETS

2.1. Land Use Map

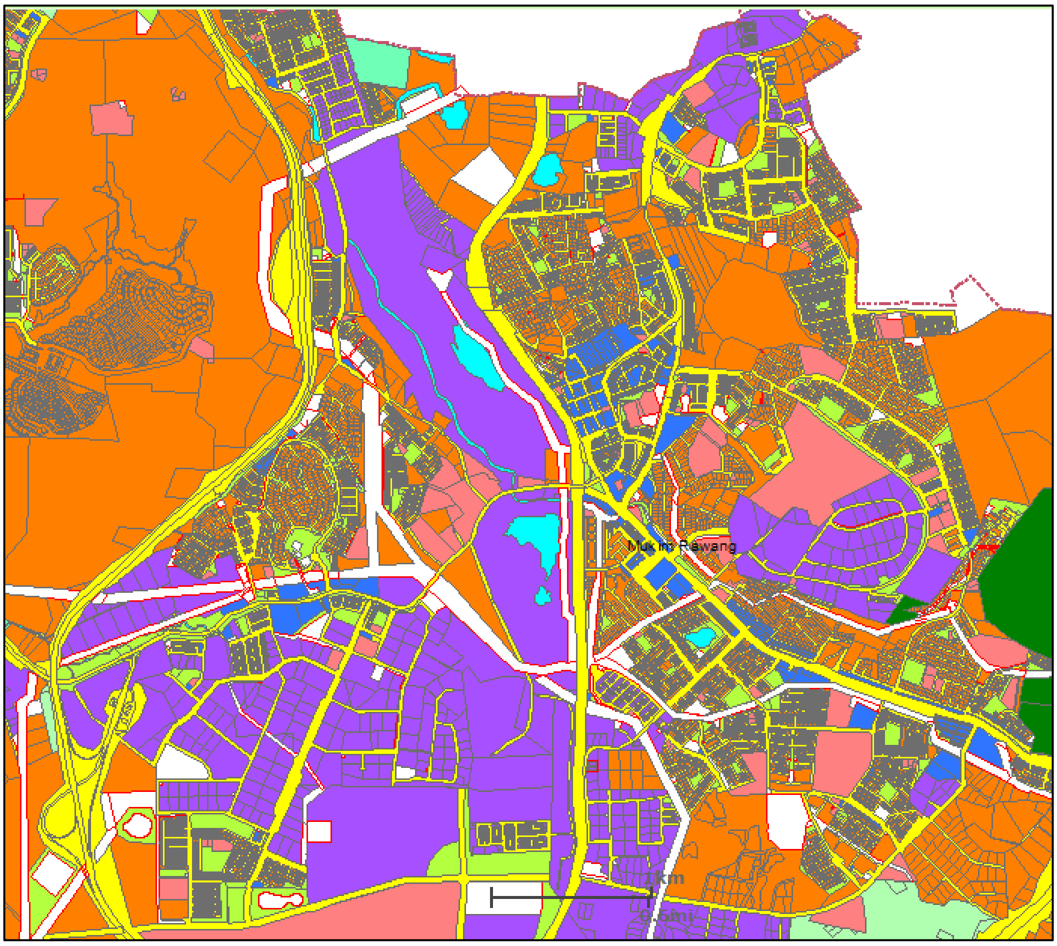

Figure 11 Land Use Map for Rawang

Table 7 Type of Land Use

Colour	Land Use	Colour	Land Use
	Residential		Infrastructure and Utilities
	Commercial		Agriculture
	Industrial		Forestry
	Institution and Community Centre		Water Bodies
	Open Space and Recreational Area		Transportation

2.1.1. Description of Land Use Data

The Department of Town and Regional Planning (DTRPS) of the State of Selangor has developed a system to monitor information on the existing and planned developments within the state digitally. The Department has applied GIS to develop a web-based system which allows the public to view the land use database for the state. This has produced a large-scale database of land use that can be used by planners or even viewed by the public on their website in map form. This database includes all 12 local authorities within the state using the broad categories of land use listed in Table 1 (including the area of residential land use – and giving the name of each neighbourhood).

These data can be purchased either in hardcopy or digital format from the Department's headquarter in Shah Alam or from a branch in each local authority. In our case, the paper map version was bought from the Department's branch in the Selayang Local Authority, Selangor.

This map shows clearly the boundary of each land use as displayed in Figure 2 so residential areas can be clearly distinguished on this map. This is important in identifying the residential areas as the first step in trying to specify the location of demand more accurately. Other types of land use such as open space can also be identified by using this map, which helps in identifying candidate locations for new schools.

2.1.2. Data Pre-Processing

The paper map had no coordinate system and had no projection system associated with it. The map was, therefore, georeferenced into the same coordinate and projection system as the other data that is the WGS 1984 system and was then used as a base map for this study. It was vectorized by digitizing the boundary lines of the land use to create polygons representing the ten types of land use employed.

This digitizing process often creates many topological errors such as slivers and duplicated lines. The data are, therefore, validated after the digitizing process is done with each of the boundaries being checked to ensure that no adjacent polygons overlap. Any duplication or slivers are removed to ensure that a sound topology can be built for the dataset.

2.2. Administrative Boundaries

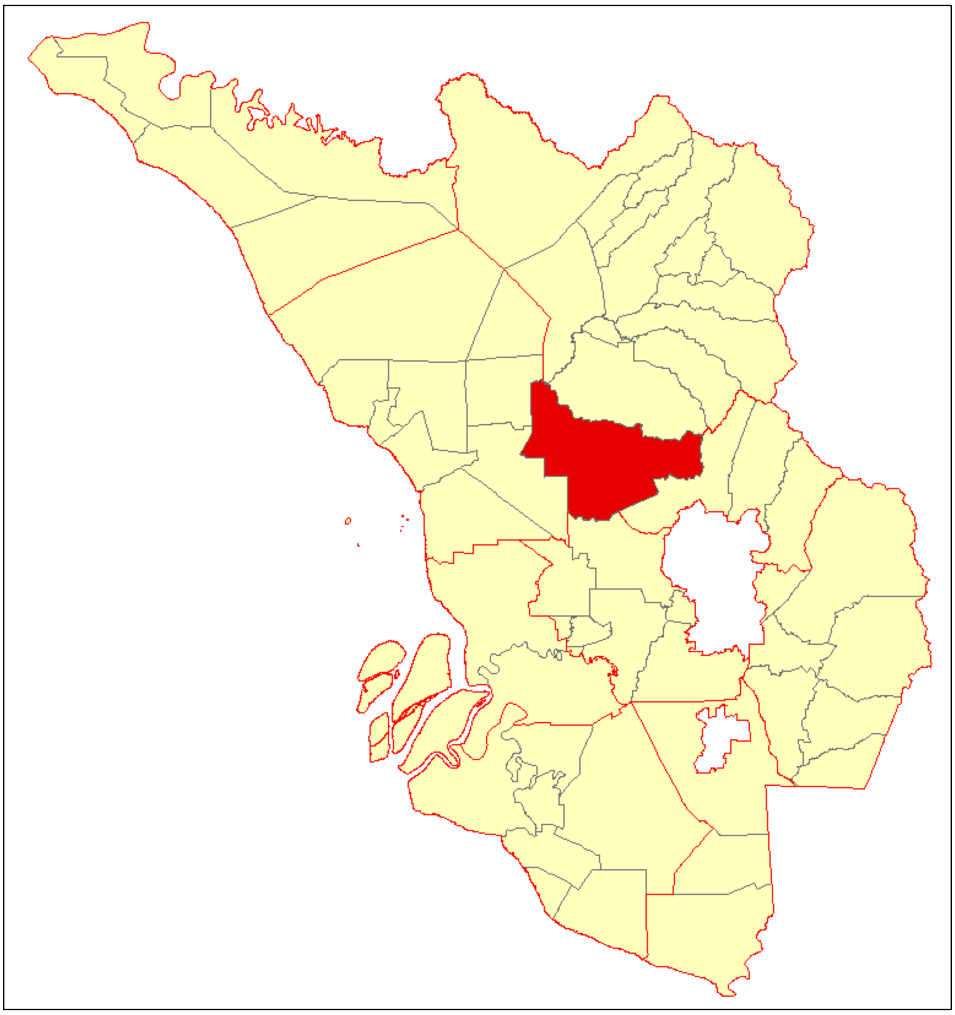

Figure 12 Administrative Boundaries within Selangor State (Red Fill shows Rawang Municipality)

2.2.1.Description of Boundaries Data

The Federal Department of Surveying and Mapping (JUPEM) has developed a fully electronic database for the whole country. This database includes all the administrative boundaries within the country from the national boundary down to state boundaries, district boundaries and municipality boundaries which show the territories of the various authorities that administer the country at different levels. For example, the state boundaries separate the country into fourteen different state authorities. Each state is then separated into a number of districts with certain responsibilities. Lastly, the boundaries of municipalities separate each district into a set of smaller local authorities. These data are part of JUPEM's restricted versions of topographic maps with a scale of 1:50000 covering the State of Selangor (Figure 3).

The data that will be used here are at the municipal level which involves a local authority managing its own area. The study area itself actually consists of part of the territory of Rawang Municipality. The municipality of Rawang lies in Gombak District and is one of the nine districts in the state of Selangor. This administrative boundary is important in defining the extent of the study. More specifically, the study area for the present research is initially defined by the area within Rawang for which detailed road network data showing minor roads are available. Essentially, this is the main built up urban part of Rawang Municipality. It is also the part of Rawang Municipality for which fairly complete cadastral data are available. However, the precise definition of the research area in fact within this built up area was finally defined by the 23 neighbourhood units for which census data was obtained.

2.2.2. Data Preprocessing

i. Data Extent

   The administrative boundary that is needed in this study is that for the municipality of Rawang only. Therefore, all other polygons were removed from the dataset.

ii. Data Validation (Topological Checking)

   - Polygon Overlap

     Each of the boundaries was checked to ensure that no adjacent polygons overlapped. This is important in developing datasets that are unique and unambiguous.

   - Line Duplication

     Thorough topological checking was carried out to ensure that there were unique boundary lines for all the polygons. Any duplication was removed to ensure that a sound topology could be built for the dataset.

2.3. Plan of Cadastral Lots for Rawang Municipality

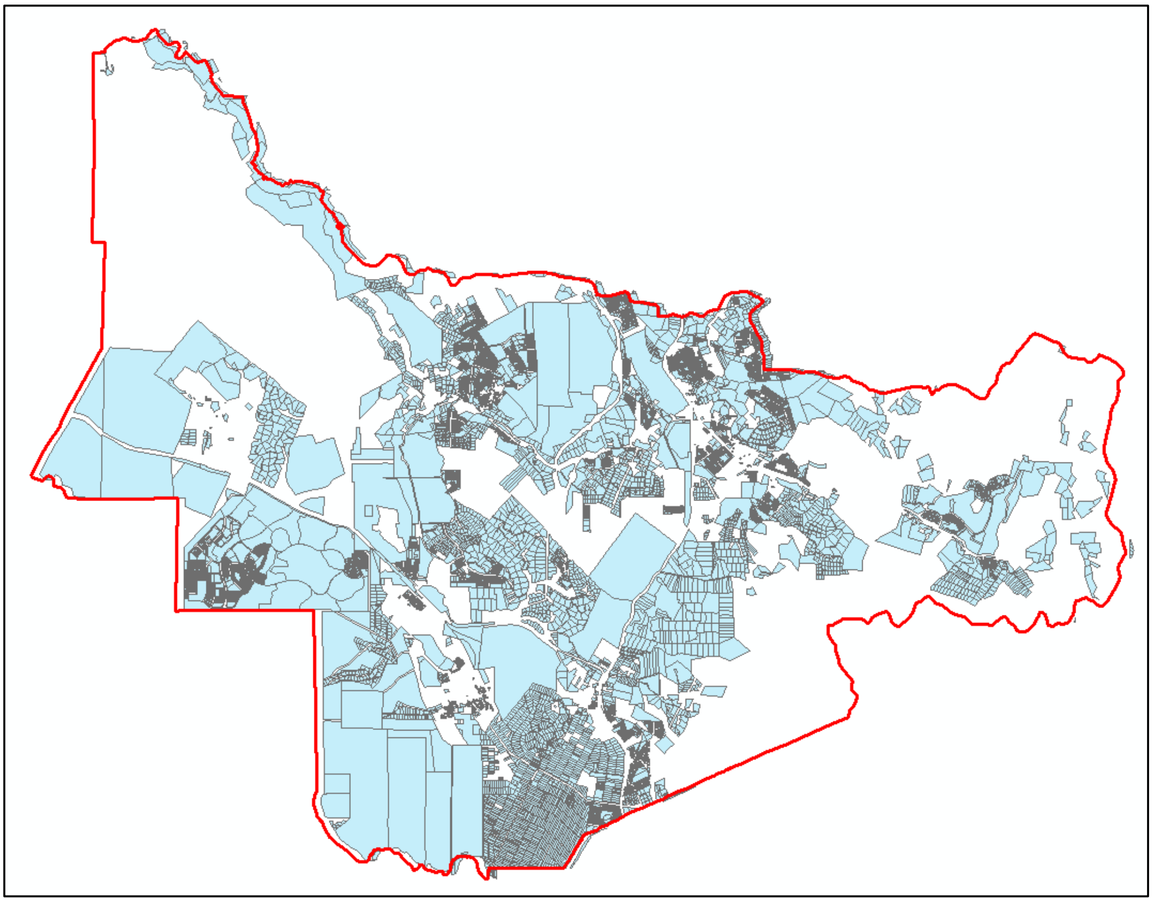

Figure 13 Cadastral Lots inside Rawang Municipality

2.3.1.Description of Data for Cadastral Lots

The digital plan of cadastral lots is derived from the National Digital Cadastral Database (NDCDB) which holds data on all ground that has been surveyed. Data for all the land that has been surveyed by the district surveyors is inserted into this database so that it is available to be viewed by the whole JUPEM community in the state of Selangor.

The data contains the boundary lines, boundary marks, polygons for the lots and any text or annotations describing the lots. These data are in the Cassini Geocentric projection system. Only the polygons for lots are needed in our research to help in specifying more precisely the locations for demand units in geographical space.

2.3.2. Data Preprocessing

i. Conversion to shapefile

The raw data of the cadastral lot plan obtained were not in shapefile format. Thus, they needed to be converted so that they could be used in an ArcGIS workstation with other vector data. These data are originally in .dxf format. Therefore, the polygons for the plot units were extracted from the datasets to create a new shapefile representing the cadastral lot units or plots.

ii. Projection and Coordinates

The NDCDB data in Malaysia is coordinated separately by each state. Each state, therefore, has its own datum system. Thus, this dataset was rectified into a new projection and coordinate system to fit the other data using the WGS 1984 coordinate system.

iii. Data Extent

The cadastral lot units needed in this study are those within the municipality of Rawang only. Therefore, all other polygons located outside the area within the administrative boundary of Rawang were removed from the dataset by overlaying the administrative polygon layer and the cadastral lot layer. The intersection of both data sets then gave only the cadastral lot units within the Rawang administrative boundary (Figure 4).

iv. Data Validation (Topological Checking)

- Polygon Overlap

  Again, each of the boundaries was checked to ensure that no adjacent polygons overlapped. This was important in establishing a dataset that was unique and unambiguous, particularly because demand data was to be allocated to these cadastral units.

- Line Duplication

  Thorough topological checking was carried out to ensure that there were unique boundary lines for these polygons. Any duplication was removed to ensure that a sound topology could be built for the dataset.

2.4. Data for the Road Network

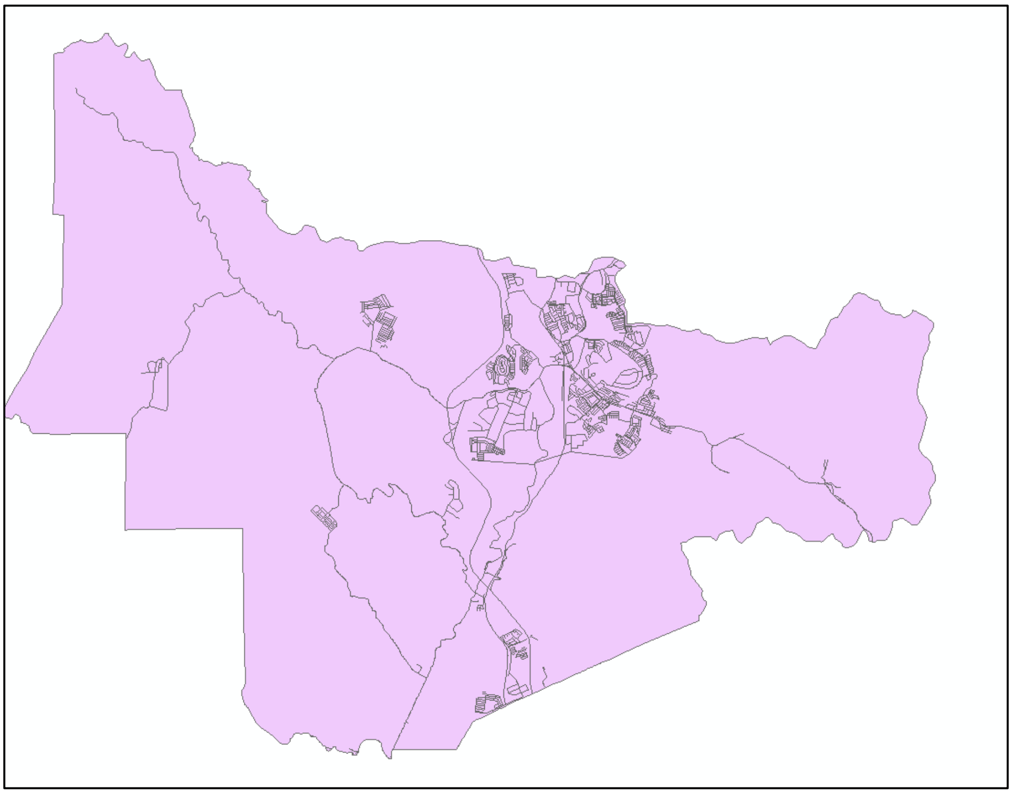

Figure 14 Road Network in the Rawang Area

2.4.1. Description of Data for the Road Network

The road network data contain four types of road which are federal roads, highways, major roads and minor roads. The federal roads and highways are built to connect major towns and cities. However, the only difference between these types of road is that highways have toll collecting fees for using them. Major roads are the main roads inside the cities and towns and minor roads are the roads inside the areas of settlements, for example the roads within residential areas.

All of these roads are included in building the network dataset. The road network data is used to find the best routes connecting pupils and schools and the distances travelled to the schools allocated.

2.4.2. Data Preprocessing

i. Data Extent

The road network used in this study is that for Rawang Municipality. Initially, all roads that are located outside the boundary of Rawang are removed from the dataset by overlaying the administrative polygon layer and the road network layer. The intersection of both data sets then gave the road network located within the administrative boundary of Rawang (Figure 5). However, as noted earlier, the study area is actually only the main built up area of Rawang town itself. This is also the area for which good detailed information on minor roads is available (the road network data obtained mainly gives main roads outside the latter area and generally gives less detail on minor roads outside this built up area).

ii. Data Validation (Topological Checking)

- Network Connectivity

  The entire network needs to be connected and snapped together at the junctions to allow turns to be done during the analysis process. This is done by tracing through the network. Any unconnected lines (i.e. links) will be snapped to closely adjacent lines by moving one of the end points of the line to the nearest line in the network to form a new node where the lines now intersect.

- Line Intersection

  The tracing process should also check the junction at the intersections of the roads. A junction should never be at lines which cross. In other words, where there are bridges over or underpass roads below another road, there should be no interpretation or assumption that the crossing point between these roads is a junction or node.

2.5. Population of Rawang by Age and Ethnicity

2.5.1. Description of Data

The population data used are from the 2010 census data collected by the Department of Statistics Malaysia (DOSM). The smallest units used by DOSM for collecting data on population and, potentially, for disseminating these data are Enumeration Blocks which typically consist of 80-120 households or 500-600 people. Despite several requests to DOSM for data at Enumeration Block (EB) level, due to confidentiality, apparently, DOSM did not provide data at EB level within the required time scale. Census data were obtained, however, for 23 neighbourhoods, created when the author divided Rawang into 23 units which corresponded as far as possible to fairly distinct residential districts. Thus, these neighbourhood units were of the author's own devising and coarser spatially than EBs and had an average size of 1670 people.

In defining these neighbourhoods so that they comprised distinct residential areas as far as possible, the author made use of boundaries between land use zones by using the land use map as discussed earlier and Google Map. Somewhat surprisingly, DOSM eventually offered to provide census data for a set of some 70 areas of Rawang, devised again by the author, and forwarded to DOSM in early May in a further attempt to obtain data at a finer spatial scale when no response had been received to an earlier request for EB data. However, this offer by DOSM to provide data at an even finer spatial scale than EBs arrived after much of the analysis had been completed using the 23 neighbourhoods as the starting point for estimating demand and was therefore much too late to incorporate such data in the research.

DOSM provides population data and data on living quarters with population broken down by gender, age and ethnicity for the defined areas as shown in Tables 2, 3 and 4. The population is divided into sixteen age groups from 0 to 75+. Ethnicity is first divided into two groups, citizens and non-citizens. Citizens are then categorized into two classes, Bumiputera and non-Bumiputera. According to the Federal Constitution of Malaysia, Bumiputera are people whose ancestors originated from the area of land (Now forming Penisular and East Malaysia). There are two categories of Bumiputera, which are Malays who speak Malay, embrace Islam and practice Malay customs and other Bumiputera who are other indigenous groups. The non-Bumiputera are immigrants who come from outside or whose forebears came to the country before Independence Day in August 31, 1957. There are three groups of non-bumiputra: Chinese, Indians and others.

Only the three main ethnic groups will be considered here i.e. Malays, Indians and Chinese, because other ethnic groups in the study area only have very small and insignificant numbers as compared to these three. These three groups are the only ones that are mentioned in the Malaysia National Education Blueprint describing the nation's education system and plans for it. Furthermore, most planning policy in Malaysia only considers these three major ethnic groups in the decision making processes.

i. Population by Gender and Total Living Quarters for Neighbourhoods

Table 8 Population of Rawang by Gender, Total Living Quarters and Households for Neighbourhoods

NEIGHBOURHOOD	Population				
	Total	**Male**	**Female**	**Total LQ**	**Total HH**
TAMAN GARING PERMAI	2,258	1,154	1,104	737	523
TAMAN GARING JAYA	423	225	198	107	101
KOTA EMERALD	948	488	460	322	255
TAMAN SRI RAWANG	707	371	336	199	183
GREEN PARK	2,985	1,468	1,517	819	736
BUKIT RAWANG JAYA	524	285	239	142	132
TAMAN RAWANG PERDANA	3,470	1,775	1,695	964	860
RAWANG PUTRA	983	517	466	324	261
TAMAN RAWANG PERDANA 2	4,943	2,524	2,419	1,623	1,347
KAMPUNG KENANGA	290	136	154	59	62
KAMPUNG DATO' LEE KIM SAI	1,253	679	574	373	322
KAMPUNG SUNGAI TERENTANG	846	473	373	201	183
TAMAN BERSATU	3,811	2,062	1,749	987	877
TAMAN JATI	1,014	505	509	276	252
TAMAN TUN PERAK	2,064	1,074	990	617	496
PERUMAHAN PKNS BATU 16	1,093	541	552	316	263
TAMAN MAWAR	1,047	555	492	339	235
TAMAN KANCING JAYA	1,688	770	918	312	269
KAMPUNG MELAYU BATU 16	737	353	384	203	192
FLAT TAMAN SETIA	391	186	205	95	89
TAMAN SETIA JAYA & TAMAN SETIA	289	146	143	106	88
TAMAN PELANGI	2,624	1,581	1,043	660	580
TAMAN SRI BAYU & TAMAN TUN TEJA	4,028	2,041	1,987	1,229	984
TOTAL	**38,416**	**19,909**	**18,507**	**11,010**	**9,290**

ii. Population by Age

Table 9 Population of Rawang by Age Group

NEIGHBOURHOOD	Total	Age group															
		0 - 4	**5 - 9**	**10 - 14**	**15 - 19**	**20 - 24**	**25 - 29**	**30 - 34**	**35 - 39**	**40 - 44**	**45 - 49**	**50 - 54**	**55 - 59**	**60 - 64**	**65 - 69**	**70 - 74**	**75+**
TAMAN GARING PERMAI	2,258	201	271	277	181	157	177	179	219	174	159	102	69	43	23	9	17
TAMAN GARING JAYA	423	19	38	46	42	43	37	37	36	30	32	24	14	15	3	2	5
KOTA EMERALD	948	96	65	71	67	108	63	97	83	43	88	91	53	10	2	6	5
TAMAN SRI RAWANG	707	46	39	57	69	67	58	71	54	39	59	65	29	24	12	10	8
GREEN PARK	2,985	128	186	253	259	233	316	201	196	185	215	213	151	168	107	88	86
BUKIT RAWANG JAYA	524	30	54	44	37	55	42	62	46	50	33	30	12	17	9	1	2
TAMAN RAWANG PERDANA	3,470	195	297	346	289	286	344	359	304	279	249	158	152	78	57	45	32
RAWANG PUTRA	983	91	87	82	67	89	120	83	100	77	66	38	37	14	22	5	5
TAMAN RAWANG PERDANA 2	4,943	519	475	502	348	362	431	421	436	432	291	190	176	131	110	65	54
KAMPUNG KENANGA	290	17	27	26	30	26	14	13	24	14	25	18	20	9	6	6	15
KAMPUNG DATO' LEE KIM SAI	1,253	55	126	106	69	82	95	102	116	115	87	53	71	56	44	47	29
KAMPUNG SUNGAI TERENTANG	846	34	34	76	69	75	82	70	63	57	53	64	40	40	37	28	24
TAMAN BERSATU	3,811	217	264	351	370	373	450	327	248	239	295	254	170	122	55	38	38
TAMAN JATI	1,014	132	113	129	60	71	109	94	92	79	60	30	16	13	5	6	5
TAMAN TUN PERAK	2,064	266	248	213	142	132	200	213	214	143	121	68	34	32	15	12	11
PERUMAHAN PKNS BATU 16	1,093	106	186	142	80	69	79	75	106	83	81	37	21	12	3	8	5
TAMAN MAWAR	1,047	82	81	119	90	202	129	59	57	68	86	34	15	13	4	3	5
TAMAN KANCING JAYA	1,688	126	62	571	296	166	86	67	39	42	90	82	21	11	8	10	11
KAMPUNG MELAYU BATU 16	737	59	53	67	54	57	63	47	41	48	52	37	47	40	29	24	19
FLAT TAMAN SETIA	391	61	47	33	26	21	17	19	23	15	43	22	13	20	6	13	12
TAMAN SETIA JAYA & TAMAN SETIA	289	112	3	51	8	24	33	12	14	11	16	1	1	1	1	1	-

TAMAN PELANGI	2,624	294	276	188	190	373	378	226	207	190	99	73	21	88	4	10	7
TAMAN SRI BAYU & TAMAN TUN TEJA	4,028	397	495	509	315	304	296	288	367	333	318	213	83	49	30	14	17
Total	**38,416**	**3,283**	**3,527**	**4,259**	**3,158**	**3,375**	**3,619**	**3,122**	**3,085**	**2,746**	**2,618**	**1,897**	**1,266**	**1,006**	**592**	**451**	**412**

iii. Population by Ethnicity

Table 10 Population of Rawang by Ethnic Group

Study Area Rawang	**Total**	**Malaysian citizens**							**Non-Malaysian citizens**
		Total	**Bumiputera**			**Chinese**	**Indians**	**Others**	
			Total	**Malay**	**Other**				
TAMAN GARING PERMAI	2,258	2,209	561	550	11	436	1,212	-	49
TAMAN GARING JAYA	423	410	130	130	-	21	259	-	13
KOTA EMERALD	948	930	332	331	1	419	177	2	18
TAMAN SRI RAWANG	707	694	106	106	-	119	468	1	13
GREEN PARK	2,985	2,900	183	169	14	1,864	845	8	85
BUKIT RAWANG JAYA	524	482	141	137	4	131	210	-	42
TAMAN RAWANG PERDANA	3,470	3,155	626	562	64	1,641	876	12	315
RAWANG PUTRA	983	939	262	255	7	455	214	8	44
TAMAN RAWANG PERDANA 2	4,943	4,728	1,631	1,611	20	1,571	1,509	17	215
KAMPUNG KENANGA	290	274	272	272	-	-	-	2	16
KAMPUNG DATO' LEE KIM SAI	1,253	1,224	228	227	1	920	70	6	29
KAMPUNG SUNGAI TERENTANG	846	806	4	4	-	766	36	-	40
TAMAN BERSATU	3,811	3,512	1,763	1,757	6	990	734	25	299
TAMAN JATI	1,014	996	668	668	-	177	150	1	18

TAMAN TUN PERAK	2,064	2,004	1,234	1,217	17	468	295	7	60
PERUMAHAN PKNS BATU 16	1,093	1,048	770	758	12	144	132	2	45
TAMAN MAWAR	1,047	988	697	697	-	145	141	5	59
TAMAN KANCING JAYA	1,688	1,662	1,207	1,207	-	311	144	-	26
KAMPUNG MELAYU BATU 16	737	691	686	680	6	-	-	5	46
FLAT TAMAN SETIA	391	370	370	370	-	-	-	-	21
TAMAN SETIA JAYA & TAMAN SETIA	289	288	39	39	-	154	95	-	1
TAMAN PELANGI	2,624	2,160	1,498	1,490	8	122	508	32	464
TAMAN SRI BAYU & TAMAN TUN TEJA	4,028	3,897	2,323	2,308	15	792	775	7	131
Total	**38,416**	**36,367**	**15,731**	**15,545**	**186**	**11,646**	**8,850**	**140**	**2,049**

2.6. Number of current pupils of each national primary school in Rawang

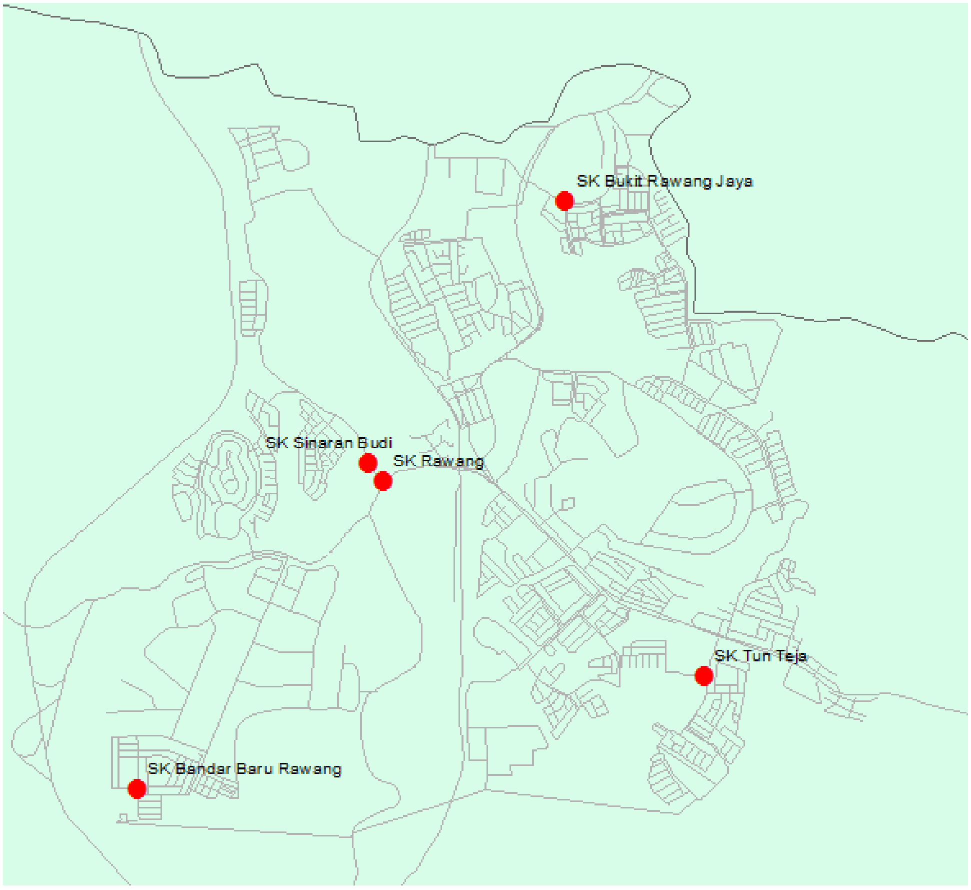

Figure 15 Location of the National Primary Schools in Rawang

2.6.1.Description of Data

Five national primary schools are located in our study area, as illustrated in Figure 6. Each school has pupils aged from 7 to 12, representing grade 1 to 6. According to MOE policy, each class can only have a maximum of thirty pupils. Two of the schools, SK Bandar Baru Rawang and SK Bukit Rawang Jaya, have a similar structure as both were built in 2000 during the same National Plans; each of them has three 4-storey buildings. The oldest national school, SK Rawang, has only two 4-storey buildings. The former English-medium school, SK Sinaran Budi, has four 3-storey buildings. The newest school, which opened in 2011, SK Tun Teja, has three 4-storey buildings. The current total numbers of pupils attending school with the number of classrooms and the maximum capacity the school can receive are displayed in the Table 5.

Table 11 Current Number of Pupils, the Number of Classrooms and the Maximum Capacity of the School

School	No. of Pupils	No. of Classrooms	Max Capacity
SK Rawang	432	16	480
SK Sinaran Budi	929	36	1080
SK Bukit Rawang Jaya	947	30	1800
SK Bandar Baru Rawang	573	30	900
SK Tun Teja	1776	47	2820

In fact, there are two kinds of school here, those which are only in session in the mornings, usually from 7.00 a.m. to 1.00 p.m., and those which have both morning and afternoon sessions (usually from, say 1.15 p.m. to 6.30 p.m.). Two of these five schools, SK Bukit Rawang Jaya and SK Tun Teja, operate with morning and afternoon sessions and, therefore, have roughly twice the capacity suggested by their number of classrooms. The other 3 schools only operate morning sessions. The capacity figures given in Table 5 reflect these differences.

3. METHODOLOGY FOR ESTIMATING DEMAND AT A FINE SPATIAL SCALE

3.1. Introduction

This chapter discusses the methodology and procedures used in more detail than was possible with the Research Paper. The first section presents a workflow chart for the methodology, Figure 7 setting out all the steps involved in the research. In the sections which follow all of the steps illustrated in the workflow chart are then explained in detail.

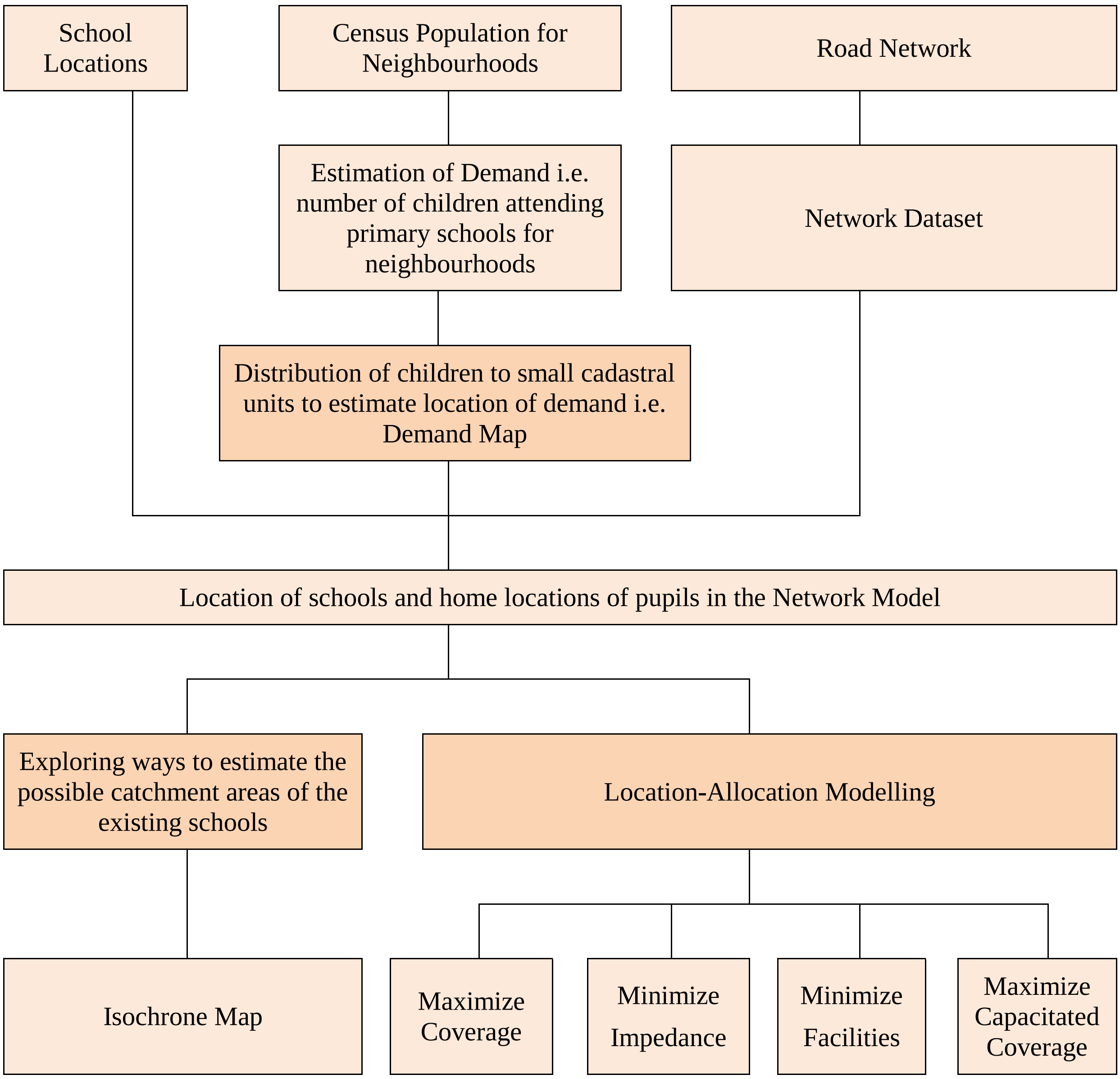

Figure 16 Workflow Chart of the Methodology

3.2. The Estimation of the Demand for Primary Schools

The demand for the service supplied by facilities is one of the key elements in LAMs. It is really important therefore to define this demand at the finest spatial scale possible to enhance the value of the solutions derived from the model. In this study, the estimates of demand are constructed from census data obtained from DOSM. As noted previously, these data provide breakdowns of population by ethnicity and age, which are essential in estimating the size of the age group which attends primary school and the proportions of each ethnic group likely to attend the national primary schools we are concerned with. The DOSM data also include the total number of living quarters in the areal units involved. These could be used later in roughly checking the location of demand at a fine spatial scale. Figure 8 shows the flowchart of the process for estimating demand for the national primary schools.

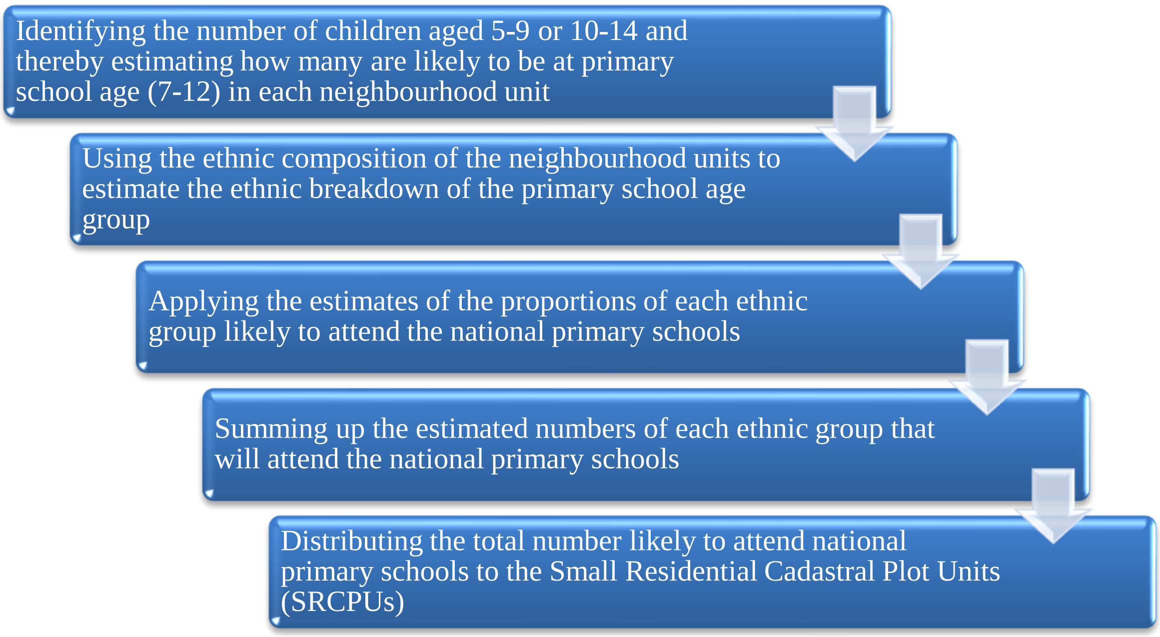

Figure 17 The Process of Estimating the Demand for National Primary School based on the Population Data

3.2.1.Mathematical Formulae for the Estimation of the Demand

The Census data groups children into three age groups relevant here with an age range of five, i.e. 0-4, 5-9 and 10-14. However, for primary school age in Malaysia is from seven to twelve years old, so the age groups defined by DOSM are not ideal. Therefore, it is necessary to first break down the latter two age groups appropriately to achieve an estimate of school-age population for each neighbourhood.

For current demand, which is the number of children aged from seven to twelve years old, the two age groups, 5-9 and 10-14 will be used. The numbers of primary school age can then be estimated by simply dividing both age groups by three fifths.

Thus,

Estimated size of (7-9) group	=	*3/5 x Population of (5-9)*

and

Estimated size of (10-12) group	=	*3/5 x Population of (10-14)*

Both of the numbers will then be added up to estimate the school-age population for each census area. This gives the potential current demand, if all children went to national primary schools. Thus

Potential Current School Age Population of each census unit	=	*Population of (7-9) + Population of (10-12)*

However, Malaysia is a multi-racial or multi-ethnic country and each state primary school in Malaysia is notionally planned to serve one of the three ethnic groups. Therefore, in estimating demand for the national primary schools, mainly intended to serve the Malay population but attended by certain proportion of Chinese and Indian children, it is also necessary to take ethnicity into consideration. In the Malaysia Education Blueprint (2013-2025), it is estimated that 90% of Malay, 8% of Chinese and 44% of Indian children attend the national primary schools (Ministry of Education (MOE), 2012). By using these percentages, the demand for these schools therefore can be estimated more accurately. First, the ethnic composition each school-age group will be estimated by using the appropriate percentage for each census unit.

Population of each Age Group by Ethnic Breakdown for the census unit	=	*Proportion of that Ethnic Group x Population of the School-age group for the census unit*
	=	*(Population of Ethnic Group/Total Population of census unit) x Population of School-age group*

Then, by taking the composition of the age group by ethnicity, and the preceding estimates made by the Malaysia Education Blueprint (2013-2025), the number of children that will attend a national primary school from that census areal unit can be estimated:

Population of Children likely to attend a national primary school for each census unit	=	*Proportion of each ethnic group attending a national primary school x Population of primary school age of that ethnic group*
	=	*(0.9 x Population of the Malay Age Group) +* *(0.08 x Population of the Chinese Age Group) +* *(0.44 x Population of the Indian Age Group)*

These values are used to estimate the demand for schools more accurately initially for the neighbourhoods.

These two sets of calculations will give the number of children that currently go to the national primary schools in each of the census units. It is however, crucial for these relatively large areal units of demand to be assigned spatially to smaller units at the finest scale possible for more precise analysis. The basic census units used were neighbourhoods. Therefore, in order to identify smaller units, the small cadastral units within each neighbourhood in the data are noted. The number of children as calculated previously for each neighbourhood is then divided evenly across all its dwelling units, whether houses or apartments within that neighbourhood. This will then give the estimated number of children per house or apartment of that neighbourhood.

Current School-age Population per Dwelling Unit	=	*Current School Age Population /* *Number of Dwelling Units*

Similarly, the likely future demand can be estimated too. The age group 0-4 would automatically be taken into account in this estimation. Next would be the population of age 5-6 years old that belonged to the age group 5-9. To obtain the population of children aged five and six, the total population of the latter age group would be split into two fifths.

Estimated size of (5-6) = *2/5 x Population of (5-9)*

Future School Age Population = *Population of (0-4) + Population of (5-6)*

Therefore;

Future School-age Population per Dwelling Unit = *Future School Age Population / Number of Dwelling Units*

This number will later be used distribute children to the map of cadastral plot units to create a spatially detailed map of demand. The procedure involved for allocating demand units to the network is explained in the next section.

3.3. Building a Network Model

The network model was built using the road data, the school locations and weighted demand points. This process of constructing the network's connectivity and attribute information for the participating feature classes in the network dataset is crucial for further analysis. Basically, the datasets are related through the road they are located on. The topological relationship of point-on-line features is used to ensure each small residential cadastral plot unit and each school is located accurately on the road network.

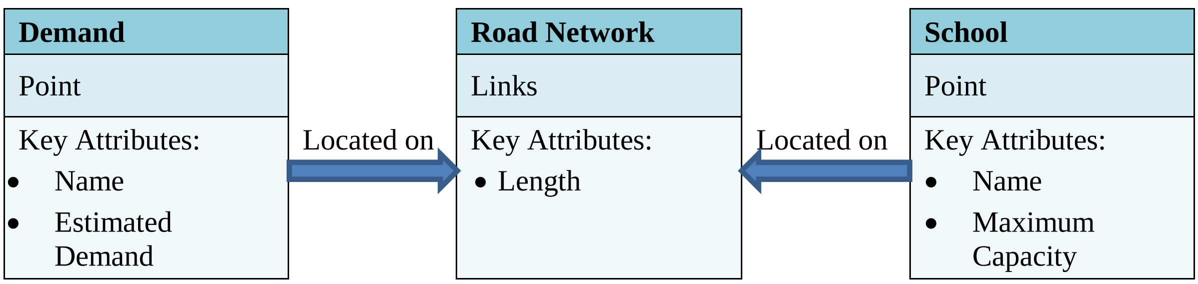

Figure 18 The Structure of the Network Model

3.3.1. Creating the Network Dataset for the Road Network

The network dataset is created by using the road network data with the cost of traversing each link specified as the distance between the two nodes at either end. A junction on the road network which includes any intersection of links is designated a node in the network dataset. These nodes, as noted previously, will allow any type of turn to be done during the later analysis.

3.3.2. Allocating Facilities (i.e. Schools) and Demand to the Network Model

i. School

The locations of schools are treated as points in the network model, both in spatial and attribute terms. Schools are located in the network by referring to their coordinates obtained from Google Maps. Since the coordinate system for the network model is the same as for Google Maps, which is WGS 1984 the schools can be located in the model by using those coordinates.

Each of the points for schools also has attribute data with the name, current number of pupils and the maximum capacity of the school. These data are important, of course, for understanding how many pupils a school can actually receive. The maximum capacity of the school is therefore a limit on the number of the pupils that can attend the school, when using a LAM with capacity constraints.

ii. Allocating Demand to Small Residential Cadastral Plot Units (SRCPUs)

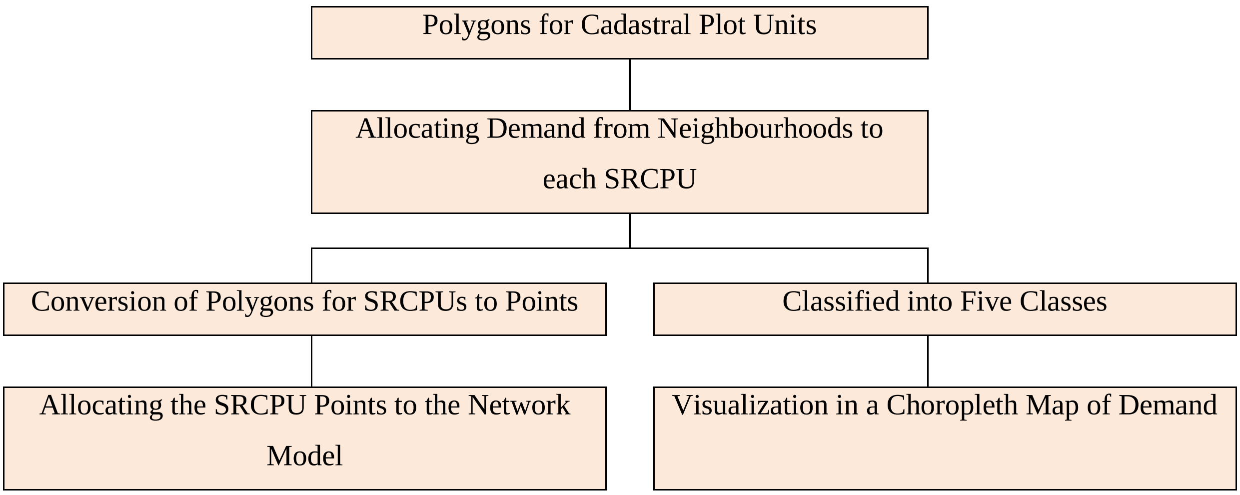

Figure 19 The Allocation of Demand for Neighbourhoods to SRCPUs

The demand for national primary school that was estimated previously at the neighbourhood level is now distributed to very small units consisting of cadastral lots (or plots) and located on the network model. The original cadastral lot units mostly represent the property boundaries of houses, factories, office buildings or other units of land ownership. As explained earlier, a map of land use types was previously used to filter out cadastral plots which are non-residential. Thus the remaining small cadastral plot units left after filtering give a very useful indication of the location of houses or apartments occupied by families. Considering that each of these dwelling units will have a possibility of pupils attending a national primary school, the estimated total demand of the neighbourhood is then assigned evenly to the dwelling units within that neighbourhood using the average number of estimated pupils per dwelling unit for that neighbourhood previously calculated and allowing for the fact that SRCPUs consisting of high rise blocks of flats will contain more dwelling units. The number of estimated pupils assigned to SRCPUs occupied by high rise blocks took account of the number of dwelling units involved but treated all as if they were located at the centre if the SRCPU involved. Therefore, the demand is assigned spatially to the whole block of the high rise building which is represented by one rather large cadastral lot unit. This is probably reasonably accurate spatially due to the fact that each of the housing units in a high-rise building will probably have a similar distance to any school they could be assigned to.

The demand is then split into five classes to visualize demand better using a Choropleth Map to show how demand is distributed across the study area.

3.4. School Districting

Two methods are used to create the maps of possible catchment areas for the national primary schools. Both methods are performed using ArcGIS software to show and explore how well the schools are located for serving population demand in the area. The notional territory of each school is shown by the boundary lines that are drawn in each method. These procedures can help to understand the pattern of access of population to the schools and can help to identify any poorly served areas.

3.4.1.Thiessen Polygons

Basically, each school will be regarded as the centroid of its surrounding polygon because the purpose of this method is to determine the notional territory of each school. The polygon of the school's territory will essentially consist of the SRCPUs for which it is the nearest school. The boundary lines indicating the areas that are notionally served by each of the schools are shown by drawing Thiessen Polygons. These lines will thus divide the study area into the five notional territories of the schools.

As mentioned previously, this method uses one of the tools in ArcGIS i.e. a tool named similarly to the method i.e. Thiessen Polygon, located under the Analysis Toolbox. This method uses Euclidean distances to generate the polygons from sets of defined centroids. The process of generating Thiessen Polygons is illustrated in Figure 11.

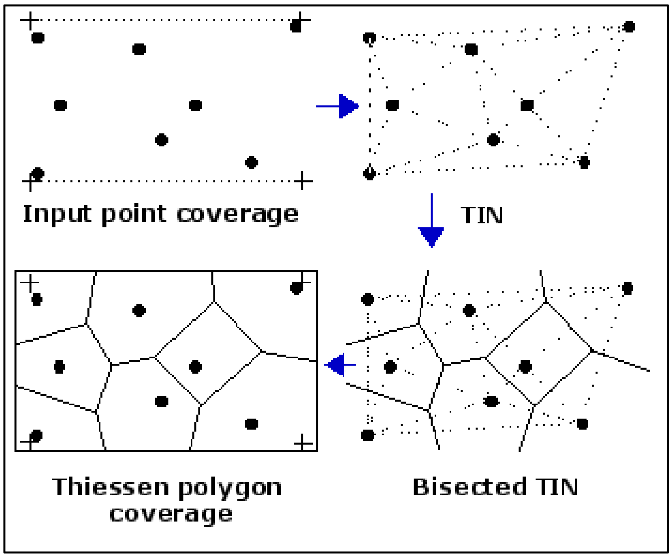

Figure 20 Process of Creating Thiessen Polygons

Technically, the polygons are created to generate a polygon topology. All of the defined centroids are first triangulated into a triangulated irregular network (TIN) and each of the triangle edges generated is then responsible for forming the Thiessen Polygons through the bisectors of the sides of each triangle being drawn. The intersections of these bisectors then determine the locations of the polygon vertices.

3.4.2.Isochrone Maps

By specifying the maximum distance from home to schools regarded as desirable by the DTRPS, the area covered within this distance by each of the schools can be mapped. This was done initially for DTRPS's ideal maximum distance of 800 meters. The map produced is essentially equivalent to an isochrone map but showing distance rather than time.

Figure 21 Steps in Creating an Isochrone Map

In ArcGIS, the tool that is used for creating Isochrone Maps is known as the Service Area Analysis tool. Located under Network Analyst, this tool uses impedance either for the distance or time travelled through the network from demand points to the central facility and requires the network dataset to be created first. Analysis can then be performed exploring the effect of using a range of covering distance or 'isochrones'. This analysis uses Dijkstra's algorithm for finding the shortest route between any two points in the network and can be used to create a set of Service Area polygons for the schools, according to a range of specified distances. The service areas generated can give insight into accessibility to the schools, as shaped by the road transport network.

In this study, the service areas within 800 meters of each school are created through points that are located on the network 800 meters from each school in question. Considering that some of the pupils might ride a bus or use other means of transportation than walking to school, a map of the areas within 1600 meters (double the recommended desirable distance) was also created. This map helps to give further insight into the accessibility of the national primary schools to pupils and helps to identify areas of high demand but poor access which can be taken into consideration as possible areas for new schools.

3.5. Location-Allocation Models

As discussed in Section 3 of the Research Paper, there are four models that are used here to solve location-allocation problems for schools, namely the Maximizing Coverage, the Minimizing Impedance, the Minimizing Facilities and the Maximizing Capacitated Coverage LAMs. The models are all used to help evaluate the existing distribution of schools and explore possible locations for additional new schools.

LAM functionality in the ArcGIS environment is known as the Location-Allocation Analysis tool, located under the Network Analysis toolbox. This analysis applies a heuristic algorithm, the Teitz and Bart algorithm, to find solutions for the LAMs involved. As a heuristic algorithm, Teitz and Bart is not guaranteed to find the optimal solution to any problem specified, but instead progressively evaluates different sets of candidate locations so that it always improves the set of facility locations until no further improvement can be obtained. The algorithm has then converged on a solution, which should be a good solution and may even possibly be the optimal solution. Essentially, this algorithm was originally designed to solve the p-median problem i.e. find the best set of candidate locations where the sum of the weighted distances from the demand nodes to their closest facility is minimized.

Before the algorithm can be implemented, it requires the origin-destination matrix of all the shortest-paths between each of the possible facility locations and all demand points through the network which is generated using Dijkstra's algorithm. In fact, Hillsman (1984) has shown how this cost matrix can be edited to solve a wide range of LAMs with other objectives and constraints, including the seven LAM problems as described previously in the Research Paper. Viewed from another perspective, effectively the Teitz and Bart algorithm generates the randomized solutions to find a group of good solutions for the required problem. As a heuristic, it may well converge on different solution when given different initial starting points for search. The group of good solutions can be combined metaheuristically in order to refine it to find even better solutions or even the best solution for the problem. This is an iterative process, which only stops when no additional improvement is possible.

The process of LAM in the ArcGIS environment starts with the creation of the network model (ESRI, 2012). This process has already been discussed previously in Section 3.3. The network model acts as the workspace for the analysis to be performed. The network model contains the facilities (schools), demand points and the road network dataset. Various properties and parameters of these elements of the analysis can be chosen to decide how the model will be run (Figure 13).

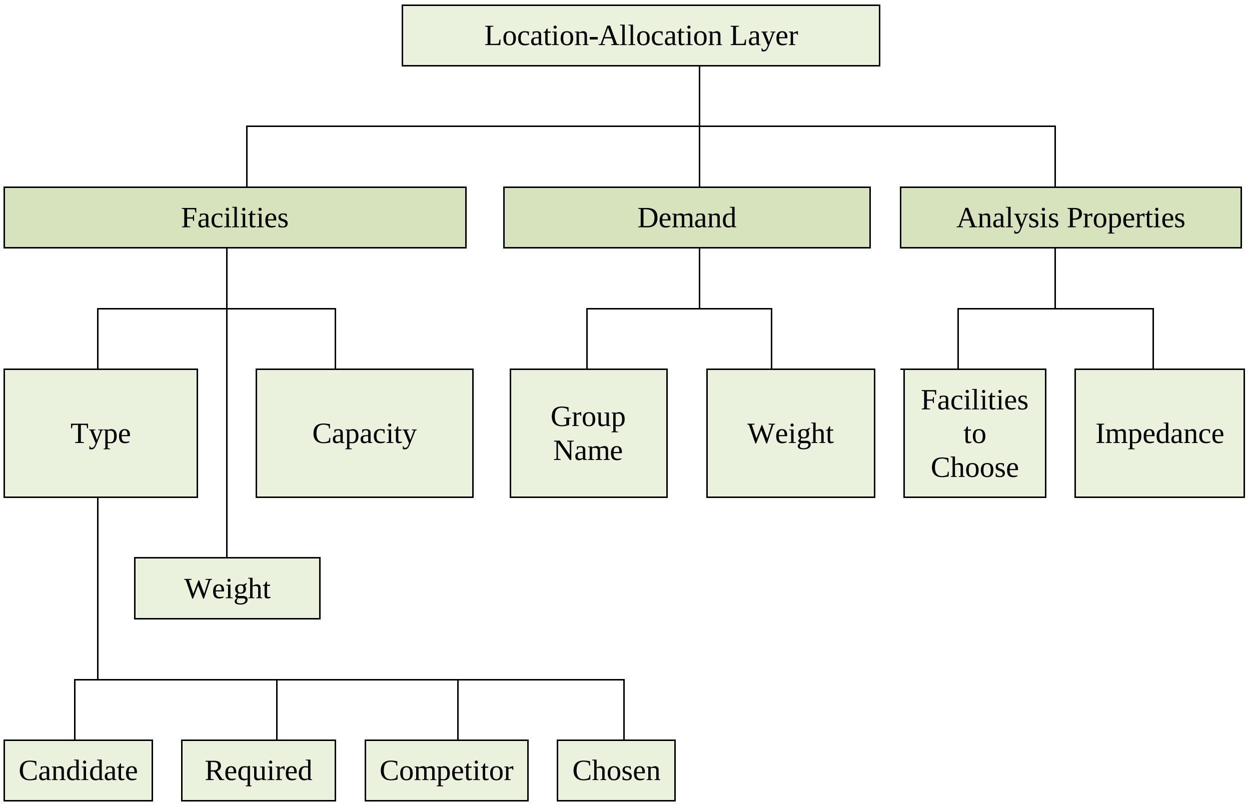

Figure 22 Procedure for Setting the LAM Layer in ArcGIS 10.1

3.5.1.Facilities

The facilities are represented by point features that hold various types of attribute or requirement for the site. It is important to specify the properties of the facilities in the analysis. There are four types of facility property, as discussed next.

i. Candidate

If a particular facility site could be part of the solution, but is not required to be in the solution, it is a Candidate facility. Normally, this type of facility is one that is tested to become part of the final solution from the analysis. The candidate facility site should be a location that is suitable for the event or service being located. There is no specific correct way to determine feasible candidate sites for facilities, which simply is a prior procedure to the LAMs. In this study, the candidate facilities are usually possible new locations for schools selected across the study area by considering various factors such as the location of poorly served areas with significant demand and sites where there appears to be space to build a new school.

ii. Required

If a facility is required to be included in the solution, it should be set as a Required facility. Considering that most of the time, planners in the Klang Valley decide to locate new schools without closing any existing schools in many runs of these models it make sense to treat the potential sites for new schools as candidate facilities and the current schools as required facilities.

iii. Competitor

A Competitor facility is a specific type of facility that is used in the target market share problem types. Such a facility represents a facility which is a rival to other facilities. For obvious reasons given earlier, this type of facility is not used here as all schools share the common function of ensuring all children will be able to attend to school i.e. they complement each other spatially rather than compete spatially.

iv. Chosen

Once the solution procedure decides that a Candidate facility is a part of the solution, it will now become a Chosen facility.

Other key properties of the facility include its weight and capacity. The weight can represent its importance or attractiveness to demand in the model. The bigger the weight of the facility, the more attractive it is to consumers in problems involving facilities competing in a market environment, as in the maximizing market share or target market share problems. In this study, since all of the schools are equally important in serving demand, the weight of each school has the default value of 1.0. The capacity property is only relevant to the Maximize Capacitated Coverage problem. Capacity will act as the limit or constraint for the amount of demand that can be allocated to the facility. When the facility reaches its maximum capacity, no further demand will be allocated, even demand within the impedance cut off.

3.5.2.Demand

A demand point is a location that represents anything that requires the services of the facilities. Usually it represents the people that need to use the facility provided. It could be a postcode centroid weighted by the actual number of people living in that area or by the estimated number. Here the demand points are cadastral lots or units in area of residential land use which can very probably be assumed to be houses or flats from which children travel to school. As with facilities, there are certain key properties of demand points that are essential for the model in allocating demand to facilities optimally.

i. Group Name (Groups of Demand Points)

   Demand points can be grouped by giving each group a distinct name. In the process of solution all members of the group may then be allocated to the same facility, if we ask for this option. However, when an impedance cut off is present, none of the group members will be allocated to a facility, if any of the group members is beyond the cut off. This property is ignored in the Maximize Capacitated Coverage, Target Market Share and Maximize Market Share problems.

ii. Weight

   Unlike facilities, the weight of demand points does not represent their importance. It is a numerical indication of the amount of demand for the service, usually the number of people requiring service at or from a facility, in this case the number of children living in a SRCPU estimated as likely to attend a national primary school.

3.5.3. Analysis Properties

Analysis parameters are set as guidelines for the model in solving the problem as specified, prior to a particular problem being run. Apart from selecting the problem type for the model to solve, there are two more key properties essential for each run.

i. Number of Facilities to Choose

   The number of facilities to be located is the most important property to be specified for any run and is a prior decision. There are no specific requirements in deciding how many facilities are to be located but normally, there should be sufficient to satisfy the demand of the region in question. This parameter has to be specified for all types of LAMs, except Minimize Facilities since the latter involves a solution process that searches for the smallest number of facilities needed to achieve a certain level of cover across the whole region.

ii. Impedance

   This property specifies the network costs attribute used to define the travelling cost along the elements in the network. An impedance cut off can be used as a constraint limiting how far search or analysis procedures will extend through the network, usually to determine the amount of demand covered within a specified distance.

Once the Network Datasets, Facilities, Demand and Analysis Properties are specified, a particular LAM can be solved. The outputs from the analysis include the facility Locations, allocation of demand points and Lines connecting demand points to the facilities they are allocated to (often called 'spider graphs'). This output can then be analysed and interpreted to understand how demand is allocated to facilities and how facilities have been selected to be in good locations to serve the demand.

4. RESULTS

4.1. Estimation of the Number of Pupils in each Neighbourhood and in Small Cadastral Plot Units (SRCPUs)

The striking areas in Figures 14 and 15 (red fill) all consist of high-rise apartment blocks The number of pupils is estimated for each of the apartments blocks instead of per dwelling unit as in the other areas with mainly one and two storey houses. Distances from all the apartment units in one block are treated as the same as one another. The estimated number of pupils for the high-rise apartments is classified into the range of 4-35 or 36-50 pupils per block. Small SRCPUs representing one and two storey houses are classified into three ranges: whether they have none, less than 1 or 1 to 3 pupils in each unit.

The most dense areas are in the south east and east of Rawang. Both of these areas mostly have at least 1 pupil per SRCPU and 50 pupils at the most where there are high rise blocks of flats. In the south east, there are 11 blocks of high-rise apartments with an average of 9 pupils per block all in the Apartment Taman Tun Teja neighbourhood. The SRCPUs in the other neighbourhoods in the south east all have 1 to 3 pupils per unit (Taman Sri Bayu, Taman Tun Teja, Taman Mawar, Taman Kancing Jaya, Perumahan PKNS Batu 17, Taman Jati, Taman Bersatu, Taman Tun Perak and Kampung Dato' Lee Kim Sai). By referring to the map of number of pupils per neighbourhood (Figure 14), it is clear that the neighbourhoods around Tun Teja are quite dense (71 to 400 pupils). These areas generally have a high percentage of Malay population and so have high proportion going to a national primary school.

The second area of high demand is in the east of Rawang. Similar to the south east, it also has several high rise apartment blocks. In fact, there are two areas of apartments there: Rawang Putra and Apartment Rawang Perdana 2 with 2 and 8 blocks of apartments respectively. These apartments are estimated to have 35 – 50 pupils per block. Other SRCPUs have 1 to 3 pupils per unit i.e. in the Rawang Perdana 1 and Rawang Perdana 2 neighbourhoods. Some of the SRCPUs, however, have less than 1 pupil per unit. This is probably due to the balanced ethnic composition of neighbourhoods in this region, mixing areas with fairly high proportions of Malay, Chinese and Indians with their varying proportions using national primary schools. Thus, the number of pupils per SRCPU varies irregularly across this area somewhat in contrast to the south east part of Rawang where the pattern is more uniform.

Table 12 Estimation of Number of Pupils in each Neighbourhood and Small Residential Cadastral Plot Unit (SRCPU)

Study Area of Rawang	Estimated No. of Pupils	No. of SRCPUs	No. of Pupils per SRCPU
TAMAN GARING PERMAI	201	1807	0.1
TAMAN GARING JAYA	30	107	0.3
KOTA EMERALD	61	330	0.2
TAMAN SRI RAWANG	32	297	0.1
GREEN PARK	178	818	0.2
BUKIT RAWANG JAYA	36	264	0.1
TAMAN RAWANG PERDANA	245	917	0.3
RAWANG PUTRA *	71	2	35.5
TAMAN RAWANG PERDANA 2	370	1243	0.3
APARTMENT TAMAN RAWANG PERDANA 2 *	399	8	50
KAMPUNG KENANGA	27	148	0.2
KAMPUNG DATO' LEE KIM SAI	108	189	0.6
KAMPUNG SUNGAI TERENTANG	49	233	0.2
TAMAN BERSATU	261	922	0.3
TAMAN JATI	116	280	0.4
TAMAN TUN PERAK	214	640	0.3
PERUMAHAN PKNS BATU 16	154	317	0.5
TAMAN MAWAR	92	330	0.3
TAMAN KANCING JAYA	315	338	0.9
KAMPUNG MELAYU BATU 16	60	302	0.2
TAMAN SETIA JAYA	1	6	0.2
TAMAN SETIA	22	165	0.1
TAMAN PELANGI	176	98	1.8
TAMAN SRI BAYU	167	198	0.8
TAMAN TUN TEJA	197	233	0.8
APARTMENT TAMAN TUN TEJA *	93	11	8.5
* High Rise Apartments with 10 floors and 10 dwelling units per floor			

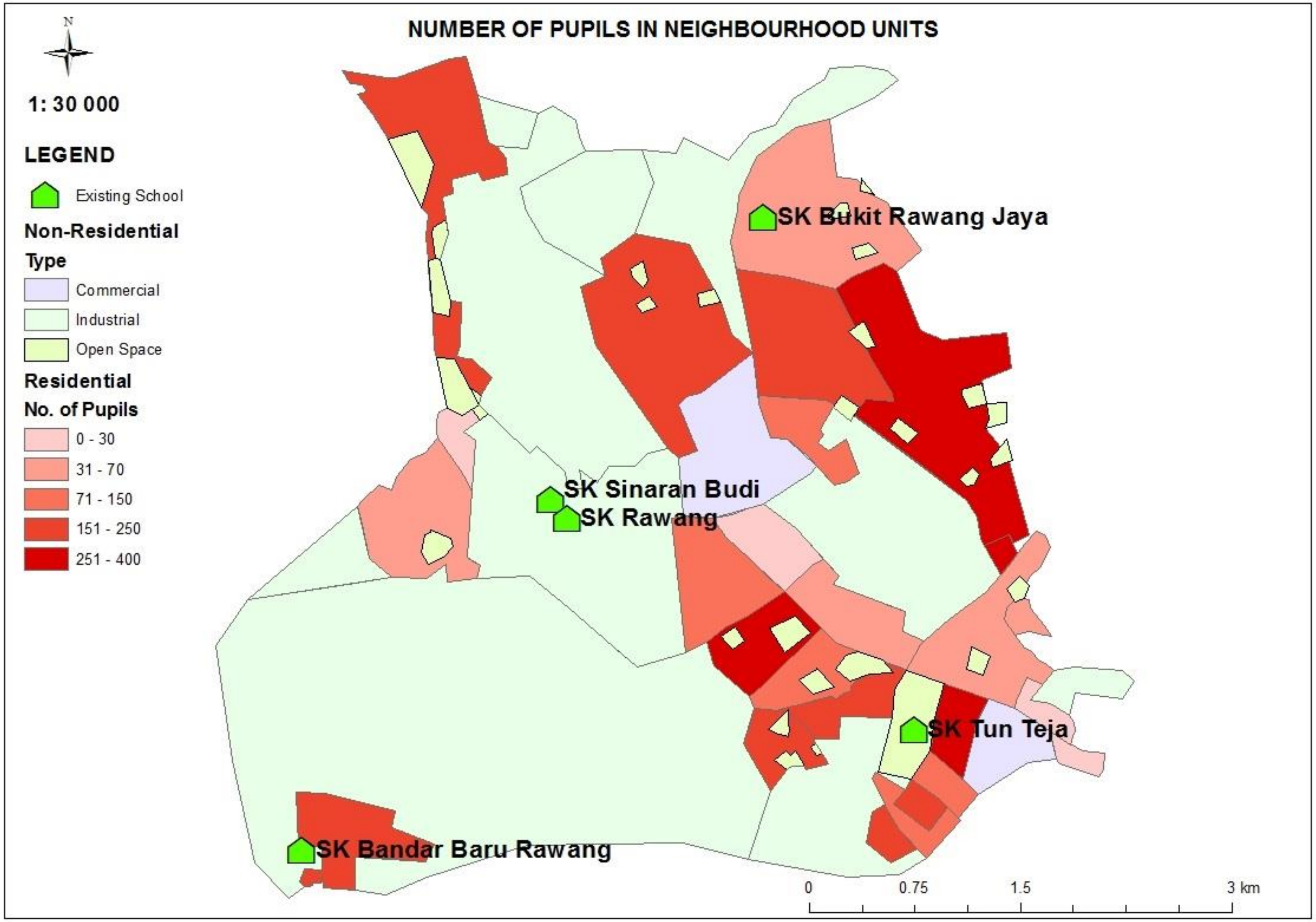

Figure 23 Estimated Number of Pupils in Neighbourhood Units

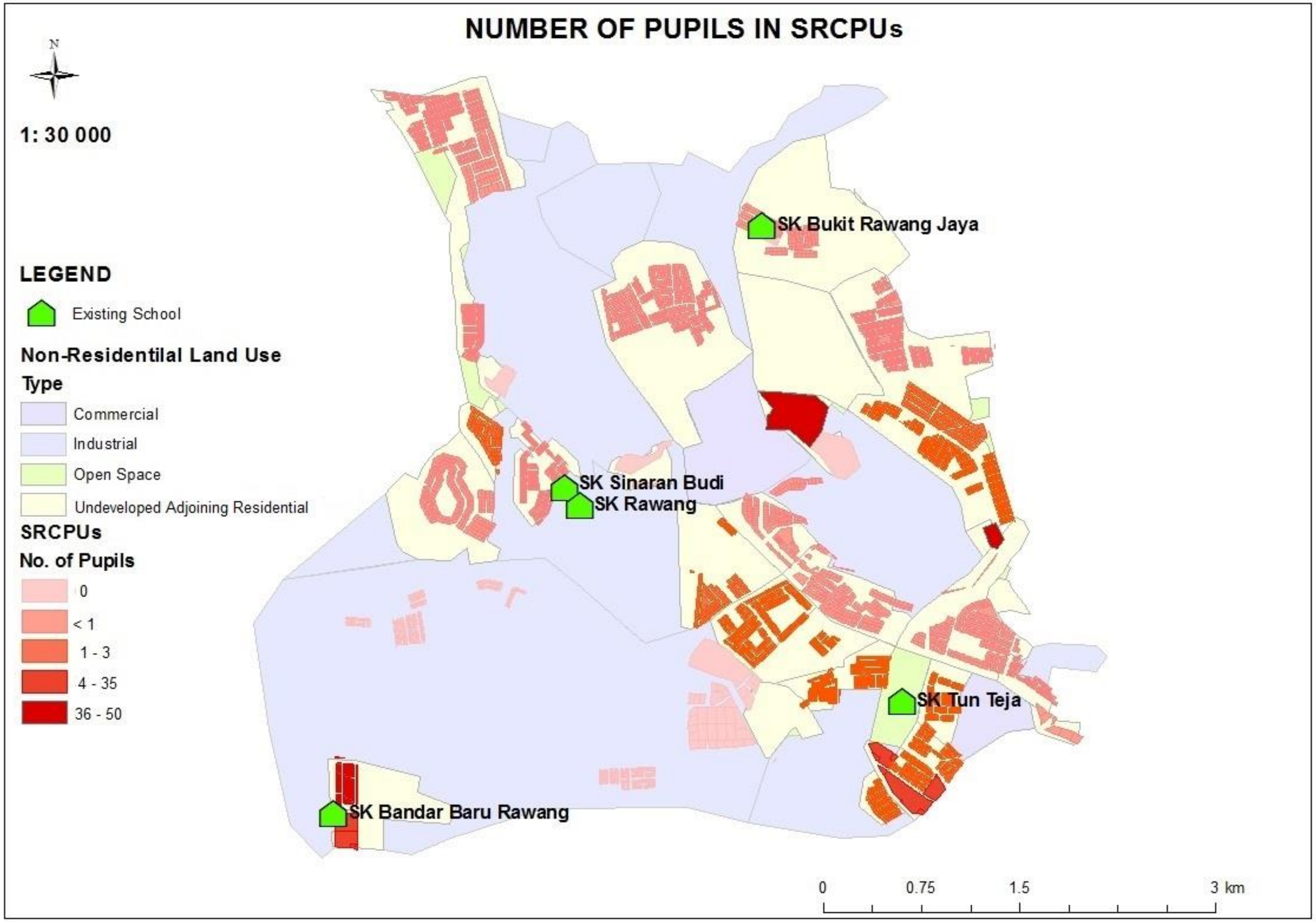

Figure 24 Demand of Schools at SRCPUs

4.2. Existing School Catchment Areas

4.2.1. Thiessen Polygons

Judging from the territorial polygons for each school, one school's polygon contains fewer residential areas than the other four. SK Rawang seems to only cover 3 neighbourhoods which have 70 pupils at the most. The SRCPUs covered by SK Rawang only have 1 pupil at the most. Furthermore, all 3 neighbourhoods are quite far from SK Rawang. Its neighbour, SK Sinaran Budi, seems to have a similarly small number of pupils (at the neighbourhood and SRCPU levels) in its territory. However, the latter seems to cover more neighbourhoods (4) which are in close proximity to each other.

SK Bandar Baru Rawang only has 1 neighbourhood within its territory (Taman Pelangi). However, this neighbourhood has quite a high number of pupils (176). The SRCPUs in its territory have an average of 1.8 pupils per unit which is the highest among any other SRCPUs in the study area without apartment blocks.

The areas in the polygons of SK Bukit Rawang Jaya and SK Tun Teja generally have high numbers of pupils. As discussed earlier, the eastern region of Rawang generally has high demand both in its south eastern and north eastern parts which is reflected in the fairly large numbers of pupils attending these two schools. Both of these schools have high capacity with morning and afternoon sessions to accommodate these high demands.

Four areas, however, are outside these 5 Thiessen polygons and are not allocated to any school's territory. One such area (Rawang Perdana 2) has quite high demand from its high-rise apartments plus 1-3 pupils per SRCPUs in its territory.

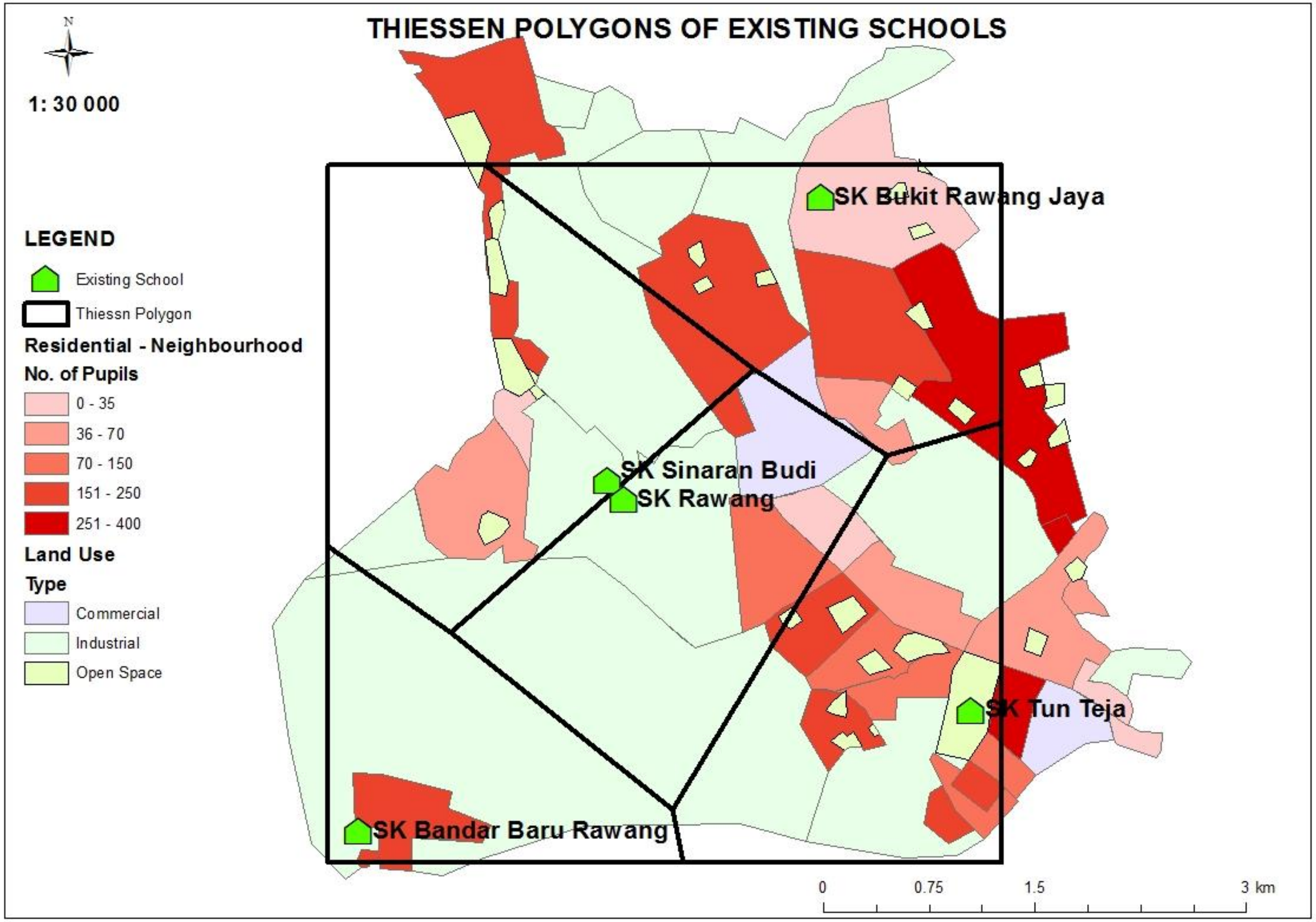

Figure 25 Thiessen Polygons for Existing Schools in Relation to Neighbourhood Areas

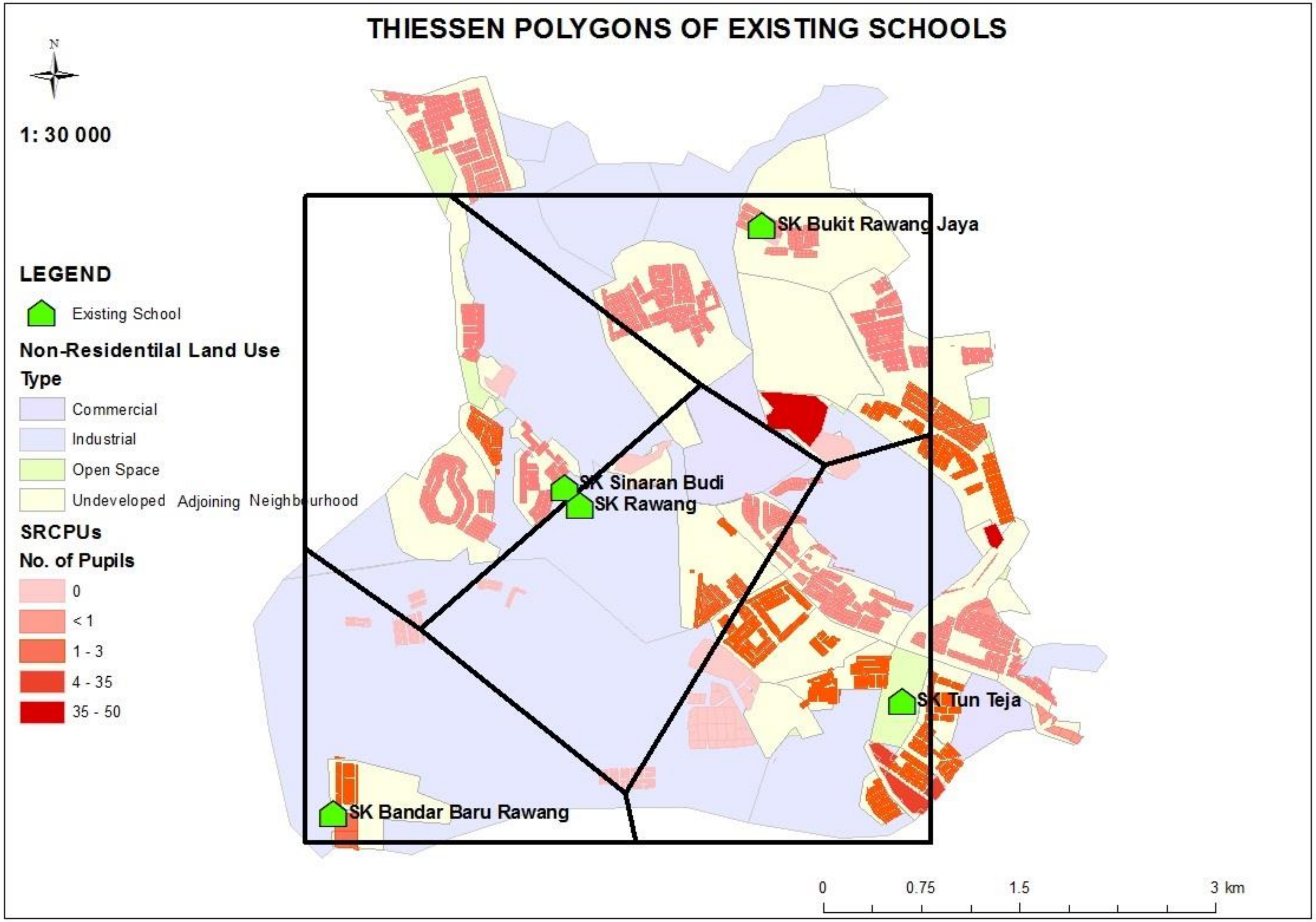

Figure 26 Thiessen Polygons for Existing Schools in Relation to SRCPUs

Table 13 Summary of Thiessen Polygons for Existing Schools

School	No. of Neighbourhoods mostly within its Territory	Remarks
SK Rawang	3	- SRCPUs with less than 1 pupil and 1-3 pupils - One all Malay area and one all Chinese area
SK Sinaran Budi	4	- SRCPUs with less than 1 pupil and 1-3 pupils - Mixed ethnic composition
SK Bukit Rawang Jaya	5	- High demand area - High rise apartments - Mostly SRCPUs with less than 1 or 1-3 pupils
SK Bandar Baru Rawang	1	- High demand area - Mostly SRCPUs with 1-3 pupils
SK Tun Teja	9	- High rise apartment areas - Mostly SRCPUs with 1-3 pupils

4.2.2. Areas with High Number of Pupils and Poor Accessibility to Schools

The eastern central area of Rawang is not covered by any school within the 800 m and 1600 m distance zones. As mentioned earlier, this area has been identified as a high demand area. According to Figures 18 and 19, it also has poor accessibility to the existing schools. Thus, this area seems to be in need of a school, probably more than any other part of Rawang.

The second area with poor accessibility is in the north west of Rawang. The closest school here is either SK Sinaran Budi or SK Rawang which are more than 3 km away. However, since it has less than 1 pupil in each of its SRCPUs, it does not have very high demand compared to the eastern area just discussed.

The area between central and south east Rawang also has quite poor accessibility to schools, though it also has quite high demand, as noted earlier. None of the schools cover this area, so it probably also has a strong case for a new school like the eastern central part of Rawang. In addition, there is only one school (SK Tun Teja) located in the adjoining south eastern area of the city to accommodate the high number of pupils living there. Thus, a new school could be needed, not only to improve accessibility, but possibly also to support SK Tun Teja in accommodating the substantial numbers of pupils in this broad sector of the city more conveniently.

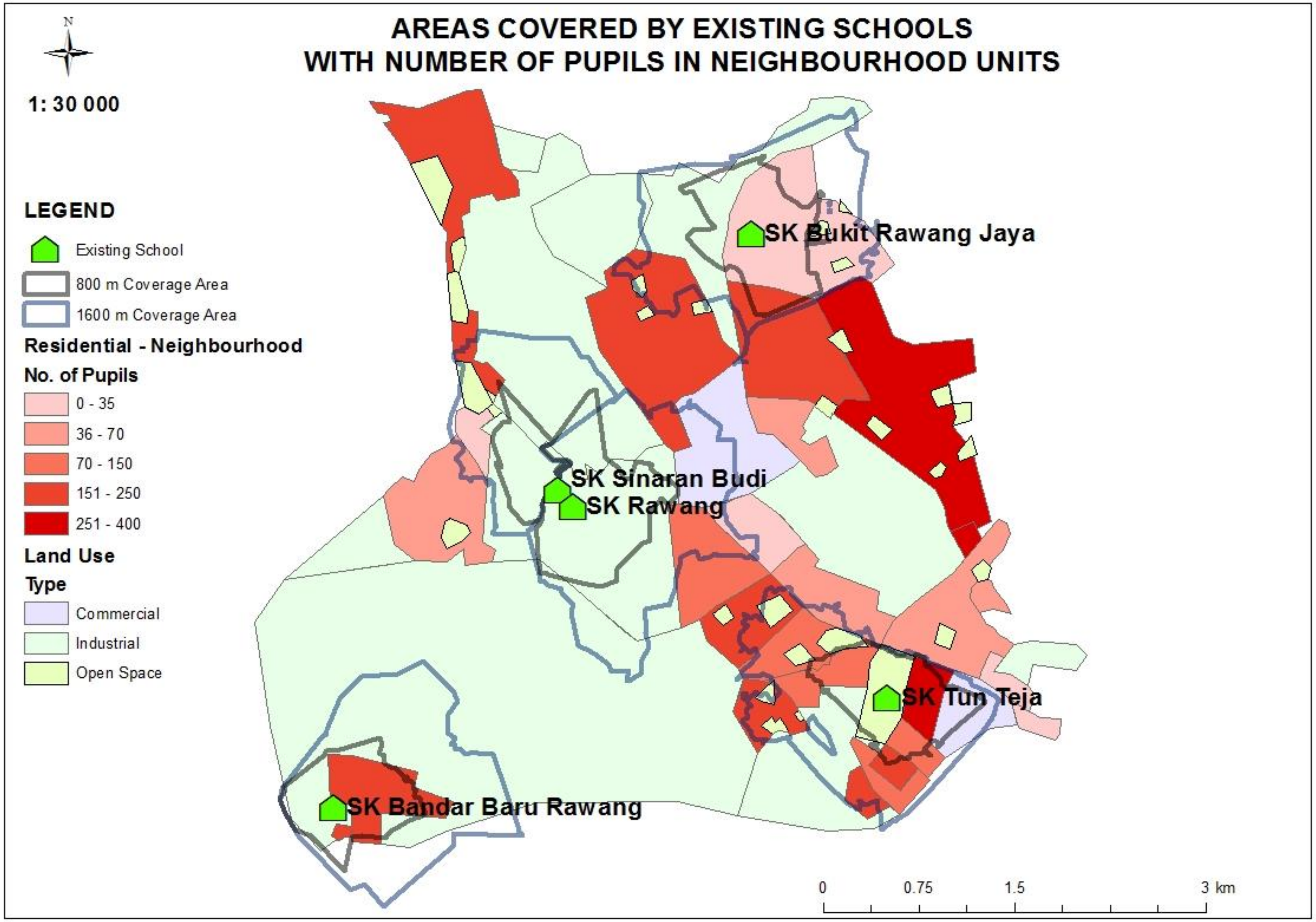

Figure 27 Areas Covered by Existing Schools in Relation to Neighbourhood Areas

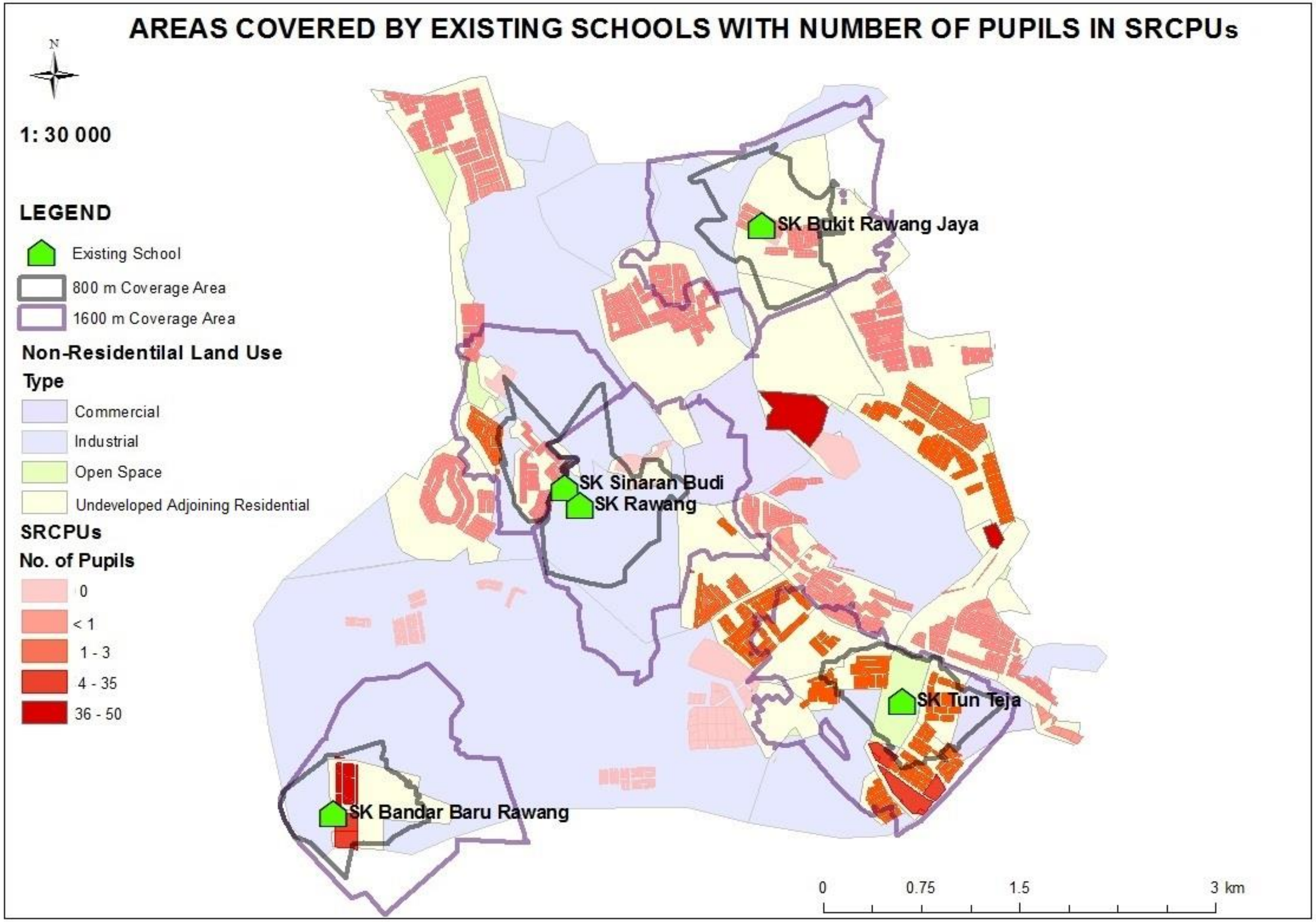

Figure 28 Areas Covered by Existing Schools in Relation to SRCPU

4.3. Candidate Locations of New Schools

Twenty seven possible candidate locations for new schools (Figures 20 and 21) were selected nearly all in high demand areas with poor access to any existing school, as defined in the previous section. The eastern central area of Rawang was prioritized as having high demand and being particularly poorly served by schools. These candidate locations were all selected in areas which seemed to have available open space for building a new school.

The north western part of Rawang was also chosen for several candidate locations since it has poor access to schools. Considering that this area is still undergoing much growth and development, it makes sense to have a school there to accommodate pupils in the future.

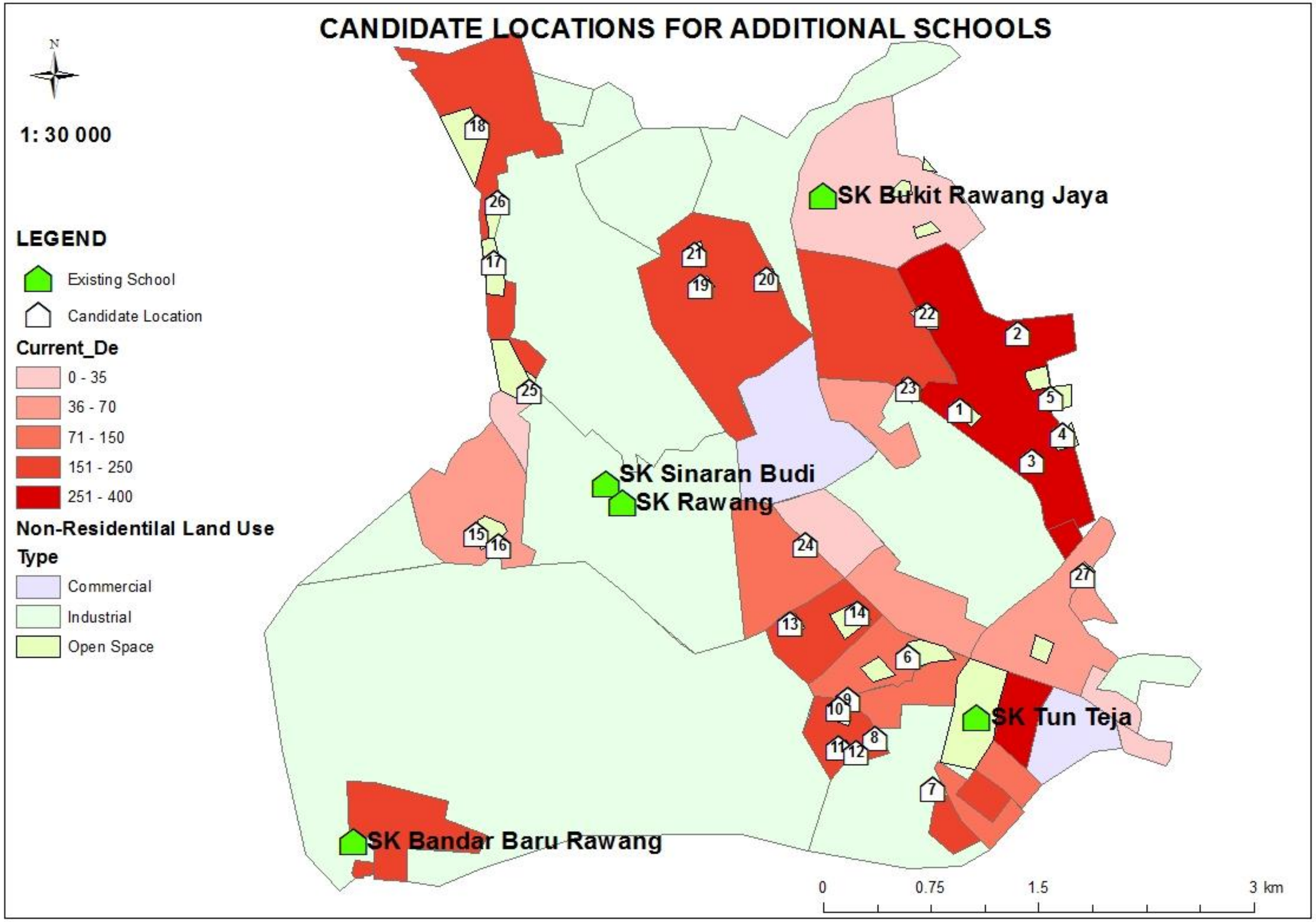

Figure 29 Candidate Locations of Schools at Neighbourhood Units

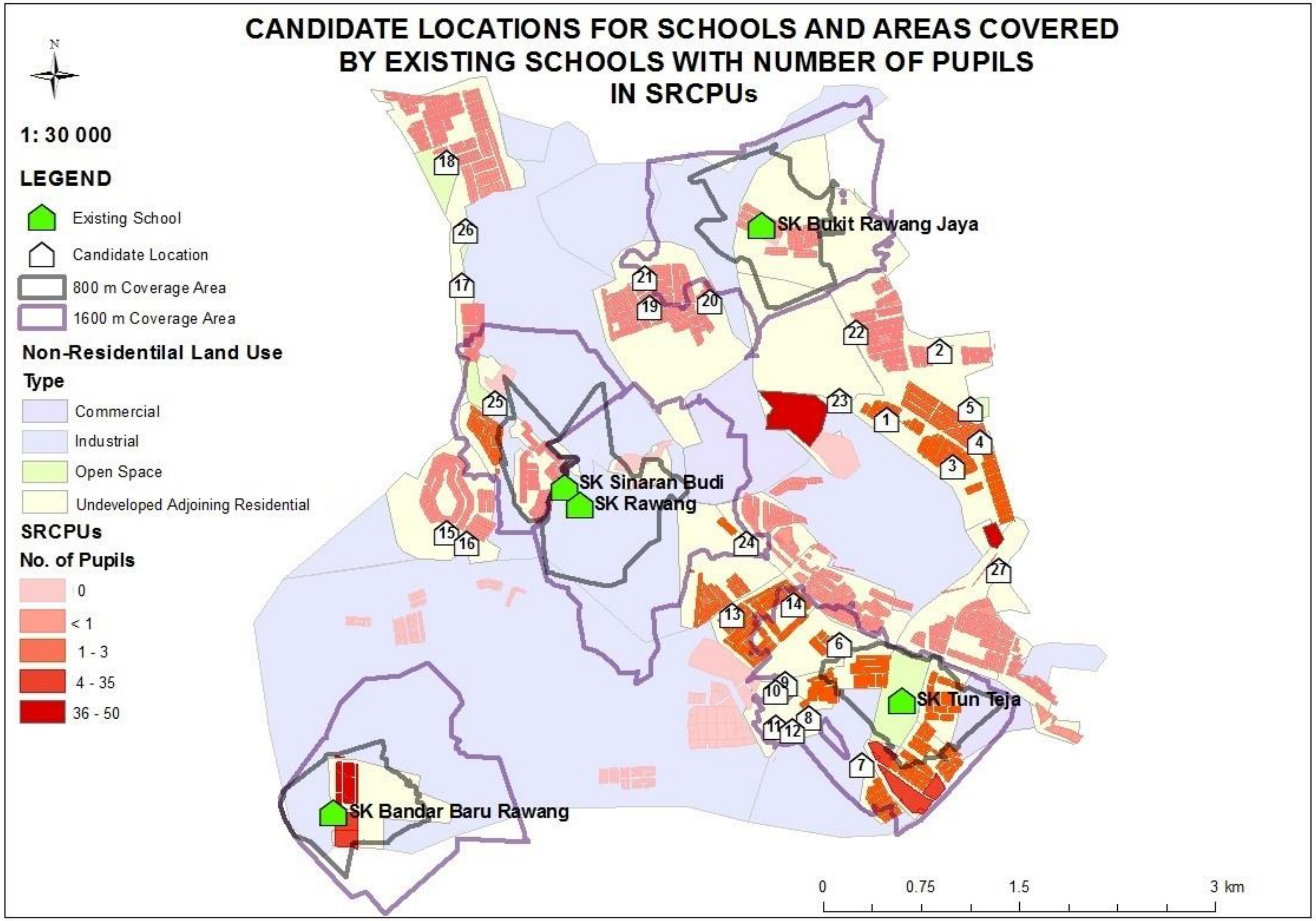

Figure 30 Candidate Locations of Schools at SRCPUs

5. CONCLUSION

The main conclusions of the research as regards the benefits of applying Location/Allocation models to the provision of national primary schools in Rawang and integrating these models with a GIS and also on how the data used in the models could be made more accurate are given in the Research Paper. The latter paper also gives the main substantive conclusions as regards good locations for siting 1 or 2 new schools to improve the accessibility of children to national primary schools in Rawang.

The present supporting document has explained in more detail than was possible in the former paper how the number of children likely to attend national primary schools was estimated by taking the age and ethnic composition of population into consideration and following the DTRPS's guidelines in planning for schools. The estimated number of pupils for 23 neighbourhoods was then distributed to a finer spatial scale of units employing 10,718 Small Residential Cadastral Plot Units to represent the houses and apartments from which children travel to school.

In choosing candidate locations for additional schools it was assumed that any sites of sufficient size in areas of green or open space were eligible. However, this selection of candidates could be made more realistic by using more detailed information on land use and land cover to eliminate areas of green or open space which are not suitable e.g. because they consist of playing fields, public parks, forest reserves, buffer zones between residential zones or other inappropriate types of green or open space.

To assess how good the 5 existing schools are at making national primary schools accessible to all the children of Rawang, two fairly simple approaches, based mainly on cartographic or visual methods, were used: Thiessen polygons and 'isochrones' based on 800 m and 1600 m travel distances. The maps produced from these methods were very helpful in identifying areas which are outside the 1600 metre coverage areas of the existing schools and can therefore be seen as badly served but also containing substantial numbers of children. The isochrones on Figures 18 and 19, especially on the latter, were particularly illuminating in this respect and gave clear suggestions as to areas meriting strong consideration if 1 or 2 more schools were to be located. Interestingly, the sites suggested from this cartographic analysis were very similar to those selected by LAMs in scenarios (c) and (d) of the research paper.

Thus the conclusions of the simple methods used here accorded well with those yielded by the more complex and sophisticated methods of location modelling of the Research paper. The latter methods, however, allow more thorough evaluation of each site in quantitative terms via criteria such as the actual numbers of pupils covered at various distances by each school and the numbers within each school's catchment. In addition, the LAMs give values for aggregate and mean travel distance within each catchment plus valuable information on the aggregate travel distance and numbers covered at various distance thresholds across the whole system. It is also very helpful that all the latter criteria can be evaluated when capacity constraints are taken into account and when such limits are ignored. Thus, while the simpler visual cartographic methods and the more complex and flexible location models can be seen as complementary to a large extent, the latter methods take the analysis to a greater depth and allow a range of potentially better locational scenarios to be assessed more broadly and more systematically through a range of criteria which can be directly compared.

REFERENCES

Abdel-Latif, A. (2007). *Combining GIS-Based Spatial Analysis and Optimization Techniques to Generate Optimum Facility Locations.* Dhahran: Saudi Aramaco.

Ahmadi Nejad Masouleh, F. (2006). *A Geographical Study of School Attendance Areas Using the Multiplicatively Weighted Voronoi Method: A Case of Rasht City, Iran.* PhD Dissertation: University of Tsukuba.

Buzai, G. (2013). Location-allocation models applied to urban public services. Spatial analysis of Primary Health Carre Centers in the city of Lujan, Argentina. *Hungarian Geographical Bulletin, 62*(4), 387-408.

Cai , Q., Rushton, G., Bhaduri, B., Bright, B., & Coleman, P. (2006). Estimating Small-Area Populations by Age and Sex using Spatial Interpolation and Statistical Inference Methods. *Transaction in GIS, 10*(577-98).

Chang, K. T. (2012). *Introduction to Geographic Information System.* New York: McGraw-Hill.

Cote, P. (n.d.). *Elements of Cartographic Style*. Retrieved 2014, from http://www.gsd.harvard.edu/gis/manual/style/

Crepper GIS. (2013). *Baltimore County Public Schools Redistricting Study 2013: Background Report.* Baltimore: Crepper GIS.

Department of Statistics Malaysia (DOSM). (2010). *Population Census Data for Rawang.* Putrajaya: Department of Statistics Malaysia (DOSM).

Department of Town and Regional Planning, Selangor. (2010). *Garis Panduan & Perancangan Piawaian Perancangan Negeri Selangor [Guidelines & Planning Policies of Selangor State]* (2nd ed.). Shah Alam: Department of Town and Regional Planning, Selangor, Malaysia.

ESRI. (2010). *Approaches to School of Business GIS Programs*. Retrieved 2014, from ESRI: http://www.esri.com/library/brochures/pdfs/approaches-to-school.pdf

ESRI. (2012). *Network Analyst*. Retrieved 2014, from ESRI: http://www.esri.com/software/arcgis/extensions/networkanalyst/key-features

Grip, R. (2013). *Redistricting Analysis for the Katonah-Lewisboro Union Free School District.* New York: Statistical Forecasting LLC.

Hakimi, S. (1964). Optimum locations of switching centers and the absolute centers and medians of a graphs. *Operations Research, 13*, 462-75.

Hakimi, S. (1965). Optimum Distribution of Switching Centres in a Communications Network and Some Related Graph-Theoretic Problems. *Operating research, 13*(462-75).

Hillsman, E. (1984). The P-Median Structure as a Unified Linear Model for Location-Allocation Analysis. *Environment and Planning, 16*, 305-318.

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2011). *Geographic Information Systems & Science* (3rd ed.). Danvers: John Wiley & Sons. Inc.

Longley, P., & Batty, M. (2003). *Advanced Spatial Analysis: the Case Book of GIS.* California: ESRI.

Lwin, K. K., & Murayama, Y. (2012). GIS Network in Geospatial Analysis. In *Progress in Geospatial Analysis* (p. 183). Japan: Springer.

Ministry of Education (MOE). (2012). *Malaysia Education Blueprint.* Putrajaya: Ministry of Education, Malaysia.

Mohammed, J., Comber, A., & Brunsdon, C. (2012). Population Estimation in Small Areas: Combining Dasymetric Mapping with Pycnophylactic Interpolation. Leicester: Department of Geography, University of Leicester.

Pearce, J. (2000). Techniques for Defining School Catchment Areas for Comparison with Census Data. *Computers, Environment and Urban Systems, 24*, 283-303.

Zandbergen, P., & Ignizio, D. (2010). Comparison of Dasymetric Mapping Techniques for Smallarea Population Estimates. *Cartography and Geographic Information Science*, *37*, 199-214.

APPENDIX

Data Sources

1. Land Use Map

Source	SISMAPS Department of Town and Regional Planning Selangor, Malaysia
Web	http://sismaps.jpbdselangor.gov.my/
Format	Paper copy
Projection	Malaysia Rectified Skew Orthomorphic
Extent	Rawang town/urban settlement area

2. Administrative Boundary

Source	Department of Surveying and Mapping Wilayah Persekutuan Kuala Lumpur dan Putrajaya
Web	https://www.jupem.gov.my/
Format	Shapefile (.shp) - Line
Projection	Malaysia Rectified Skew Orthomorphic
Coordinate System	WGS 1984
Extent	State of Selangor, Malaysia

3. Road Transportation Network

Source	Department of Surveying and Mapping Wilayah Persekutuan
Web	https://www.jupem.gov.my/
Format	Shapefile (.shp) - Line
Projection	Malaysia Rectified Skew Orthomorphic
Coordinate System	WGS 1984
Extent	State of Selangor, Malaysia

4. Cadastral Lot Map

Source	Department of Surveying and Mapping Selangor
Web	http://spak.jusl.gov.my/
Format	AutoCAD Drawing Exchange Format (.dxf)
Projection	Cassini Geocentric
Geographic Coordinate System	GDM 2000 Selangor
Extent	State of Selangor, Malaysia

5. Population by Age and Ethnicity

Source	Department of Statistics Malaysia
Web	http://www.statistics.gov.my/
Format	Spreadsheet
Extent	Rawang urban settlement area

6. Number of current pupils of each national primary schools in Rawang

Source	District Education Office of Gombak
Web	http://www.pelajarangombak.net/v1/
Format	Shapefile (.shp) - point
Projection	WGS 1984
Extent	Rawang urban settlement area